On the mechanism of promiscuous ligand-binding by multidrug resistance regulator LmrR and structural investigations of lytic transglycosylase MltF

Madoori, Pramod Kumar

On the mechanism of promiscuous ligand-binding by multidrug resistance regulator LmrR
and
Structural investigations of lytic transglycosylase MltF

The research presented in this thesis was carried out at the Protein X-ray crystallography group, Laboratory of Biophysical Chemistry, Groningen Biomolecular Science and Biotechnology Institute (Faculty of Mathematics and Natural Sciences University of Groningen) and was supported by Ubbo Emmius Bursary, University of Groningen.

RIJKSUNIVERSITEIT GRONINGEN

On the mechanism of promiscuous ligand-binding by multidrug resistance regulator LmrR and

Structural investigations of lytic transglycosylase MltF

Proefschrift

ter verkrijging van het doctoraat in de
Wiskunde en Natuurwetenschappen
aan de Rijksuniversiteit Groningen
op gezag van de
Rector Magnificus, dr. E. Sterken,
in het openbaar te verdedigen op
vrijdag 25 mei 2012
om 12:45 uur

door

Pramod Kumar Madoori

geboren op 1 juli 1975
te Nizamabad, India

Promotor:	Prof. dr. B.W. Dijkstra
Copromotor:	Dr. A.M.W.H. Thunnissen
Beoordelingscommissie:	Prof. dr. A.J.M. Driessen Prof. dr. D.J. Slotboom Prof. dr. O.P. Kuipers

ISBN: 978-90-367-5509-2 (electronic version)

ISBN: 978-90-367-5510-8 (printed version)

For my wife, daughter, my parents, brothers and my In-laws
and all my teachers
and all my friends

Table of contents

Scope of the thesis

This thesis describes structural studies of two proteins from two biological systems. The first part (Part I) consists of chapters 1, 2 and 3, while the second part (Part II) consists of chapters 4 and 5 and deals with the lytic transglycosylase MltF from *Escherichia coli*.

The first part of the thesis discusses the mechanism of multidrug recognition by the transcriptional regulator LmrR from *Lactococcus lactis*. LmrR regulates the expression of the multidrug transporter LmrCD, which belongs to ABC superfamily and is primarily responsible for the multidrug phenotype of *L. lactis*. By binding various ligands similar to its cognate transporter, LmrR undergoes a structural change facilitating the upregulation of LmrCD. Understanding the features of multidrug recognition will allow further expansion of the multidrug recognition principles, as this is the first example of a transcriptional regulator that governs a multidrug ABC transporter.

The second part of the thesis consists of chapters 4 and 5 and deals with the lytic transglycosylase MltF from *Escherichia coli*. Lytic transglycosylases (LTs) are enzymes involved in peptidoglycan (PG) metabolism. They catalyze two reactions on PG. One is the cleavage of the β-1,4-glycosidic bond, similar to lysozyme, the other is the subsequent intra-residue transglucosylation reaction, in which a 1,6-glycosidic bond is formed, thereby producing GlcNAc-anhydro-muropeptides. These end-products are the hallmark of LTs. In *E. coli* different types of LTs are present, but specific roles of them in the *E. coli* metabolism remain unclear. These enzymes differ from each other by having additional domains for which the function is often unknown. MltF differs from other LTs by having an N-terminal periplasmic binding domain (MltF-NTD) and a common C-terminal lysozyme like fold. Structural investigations on MltF may allow a better understanding of MltF-NTD.

Chapter 1 - Introduction

Why is there an urgent need to tackle antibiotic resistance in pathogenic microbes? How do microbes gain resistance to the available antibiotics? What are the principles of MD recognition? Why do we need to study MD recognition? These questions are addressed in this chapter by describing resistance mechanisms, multidrug transporters, and several reasons for the need of MD-related transcriptional regulators in pathogenic microbes. A short overview of the canonical and non-canonical types of multidrug recognition is presented to understand how LmrR multidrug recognition differs from the canonical recognition, which is the main subject of this thesis.

Chapter 2 - Structure of the transcriptional regulator LmrR and its mechanism of multidrug recognition

MD recognition principles of the members of PadR family are not known. Chapter 2 describes the structural analysis of LmrR. Multidrug binding principles are derived by comparison of structures between apo and drug bound forms of LmrR. LmrR bound to Hoechst 33342 and daunomycin emphasizes the role of a single drug binding site that can accommodate structurally different ligands. Primarily, dimer-related tryptophan residues play a dominant role in aspecific binding, in which the indole rings sandwich the planar ring system of different compounds. The observed drug binding features in LmrR differs considerably from the current views of MD recognition.

Chapter 3 - Multidrug resistance regulator LmrR uses a single site to bind structurally diverse compounds

Chapter 3 describes additional structures of LmrR in complex with two chemically different ligands, ethidium and riboflavin. These structures emphasize the role of a single ligand-binding site at the dimer interface for MD recognition.

Chapter 4 - Purification, crystallization and preliminary X-ray diffraction analysis of the lytic transglycosylase MltF from *Escherichia coli*

In this chapter the purification, crystallization and preliminary X-ray analysis of two soluble, C-terminally His_6-tagged forms of MltF lacking residues 1-22 are described. One form of the protein contains both domains (sMltF), whilst the other contains only the N-terminal domain (sMltF-NTD).

Chapter 5 - Crystal structure of the SBP-like N-terminal domain of *Escherichia coli* lytic transglycosylase MltF

Chapter 5 describes structural features of MltF-NTD. We present the structure and compare it with structural homologs and highlight that the ligand-binding region is significantly different.

CHAPTER

General introduction

1

1

Rise and fall of antibiotics efficacy

In 1929, Alexander Fleming observed that a mold of contaminating *Penicillium notatum* on a *Staphylococcus* agar plate prevented the growth of bacteria. This finding marked the beginning of the development of antibiotics to combat infectious diseases, and it was one of the most valuable discoveries in the 20th century. Nevertheless, the use of natural resources for the treatment of various infections can be dated back to over thousand years ago [1], most often referred to as "grandma recipes", some of which are still in practice. Fleming's discovery led to the identification of penicillin as the first antibiotic. It was considered a wonder drug because it appeared to be able to combat all kinds of pathogenic organisms. The discovery of penicillin was followed later by a rapid discovery of other new antibiotics, not to have a library of different entities *a priori,* but as weapons to reduce the rise of deadly infections [2]. However, pathogens that were once effectively treated with a particular antibiotic, have been gaining alterations in their physiological systems, enabling them to neutralize the toxicity of the antibiotics. This phenomenon is called resistance. Resistance can be transferred across different taxonomic groups by mobile genetic elements, like naked DNA or transposons, plasmids, and bacteriophages [3]. In the absence of mobile elements, step-wise mutations in chromosomes can also lead to resistance in bacteria [4; 5]. This process was shown to be responsible for the emergence of resistance against penicillin and tetracycline in *e.g. Neisseria gonorrhoeae*. Because of improper use of antibiotics and continuous exposure to new generations of antibiotics and because resistance can easily be transferred, microbes have developed, in a relatively short period of time, resistance to various classes of antibiotics [2; 6; 7]. Moreover, they developed resistance to multiple antibiotics, leading to the emergence of a multiple antibiotic resistance (mar) phenotype. Notable examples in this case are strains of *Mycobacterium tuberculosis*, *Enterococcus faecium*, *Enterobacter cloacae*, *Klebsiella pneumoniae*, *Staphylococcus aureus*, *Acinetobacter baumannii* and *Pseudomonas aeruginosa* [8; 9]. Rather than using a single antibiotic, mar requires the use of several antibiotics to treat a specific infection. Such a treatment, unfortunately, at the same time accelerates the evolution of new types of resistant strains [10; 11]. As a result, resistant strains are rising at a rate that is not matched by the discovery of new antibiotics.

Mechanisms of antibiotic resistance

The biological route by which microbes demonstrate resistance varies and can be divided into different categories: (*i*) direct action against the antibiotics; (*ii*) the alteration of the antibiotic target inside the cellular milieu, making the drug unable to inhibit a vital function, and (*iii*) drug induced oxidative stress. An example of a direct action of microbes against an antibiotic is the expression of enzymes, such as β-lactamases, which destroy penicillins and cephalosporins [12], and of membrane-bound pumps that actively excrete the antibiotics from the cell. Alteration of the antibiotic target inside the cellular milieu is exemplified by the synthesis of ribosomal protection proteins (RPP), which induce conformational changes in the ribosomes [13]. Ribosomes are the target of the antibiotic tetracycline, and by changing the conformation of the ribosome microbes have developed resistance against this antibiotic. In the third category, the drug itself can act as a mutagen and give rise to drug resistant strains. Certain antibiotics like β-lactams and quinolones induce the formation of reactive oxygen species (ROS), which can damage the DNA, and activate the error-prone SOS response. Thus ROS in conjunction with SOS leads to emergence of mutations, which results in the emergence of a resistant strain [14;15]. Details of resistance mechanisms have been reviewed by Alekshun & Levy [16]. Except for the efflux pumps, the resistance mechanisms described above are limited to a certain class of antibiotics. In contrast, efflux pumps have the ability to extrude a broad spectrum of chemically distinct cytotoxic molecules [17]. Organisms which possess these pumps, be it prokaryotes or eukaryotes, often show a multidrug resistance (MDR) phenotype [18; 19; 20].

Multidrug efflux pumps in bacteria

All efflux pumps/transporters reside in the cell membrane and they can be divided into two families based on bioenergetic and structural features: (*i*) primary transporters which use ATP as their source of energy, also known as ATP–binding cassette (ABC) transporters (*e.g.* sav1866 from *Staphylococcus aureus* and the mammalian P-glycoprotein) and (*ii*) secondary transporters which utilize the proton (or sodium) gradient as a source of energy. On the basis of their structural organization and substrate specificity, secondary transporters in multi-drug resistance have further been classified into four

families: (*i*) the multidrug and toxic compound extrusion (MATE) family (*e.g.* NorM from *Erwinia amylovora*); (*ii*) the major facilitator superfamily (MFS; *e.g*. EmrD from *E. coli*); (*iii*) the small multidrug resistance (SMR) family [21] (*e.g.* EmrE from *E. coli*) and (*iv*) the resistance nodulation division (RND) family (*e.g*. AcrB from *E. coli*) [22]. Structures of the multidrug transporters with bound drugs are limited. The only structures of MDR pumps with bound substrates are those of P-glycoprotein bound to cyclic peptide inhibitors (4.4 Å) [23], AcrB with minocycline and doxorubicin at 3.1 and 3.3 Å resolution [24] and the 7.0 Å structure of the complex of EmrE with TPP^+ (tetraphenylphosphonium) [25]. These complexes, though limited in number and resolution, indicate the importance of a flexible voluminous drug-binding pocket containing mainly aliphatic and aromatic hydrophobic amino acids for multi drug recognition.

Despite their role in drug resistance, efflux pumps usually have a basal level of expression because they also perform a physiological role to provide protection for the microbes in their ecological niche. For instance, enteric bacteria (which reside in the gastrointestinal tract of humans and animals) use the AcrAB pump to survive in the hydrophobic environment constituted by bile salts and fatty acids[26]. The multidrug transporter Blt of *Bacillus subtilis* is normally co-induced with enzymes involved in spermidine/spermine metabolism and its normal function is a spermidine transporter [27; 28]. The chemical character of several toxic compounds designed against pathogens is similar to the compounds normally encountered by bacteria in their natural environment. For example, many natural and synthetic cytotoxic compounds contain planar aromatic groups, have a lipophilic nature and are often weakly cationic. The basal level of expression and broad substrate specificity of MDR efflux pumps ensure that a bacterial cell is always ready to defend itself against environmental toxins[29]. However, the basal expression level of MDR pumps is not sufficient to support resistance to toxic compounds at higher concentrations and, therefore, in several cases, efflux pumps are upregulated. The upregulation is achieved by sensory molecules residing in the cell that are referred to as "transcriptional regulators", which control the expression of their cognate transporters as well as their own.

Fine tuning of the expression of MDR pumps

Considering their broad substrate specificity, MDR pumps can be potentially dangerous to the cell. For example, constitutive expression of the tetracycline resistance pumps TetA(B) and TetA(C) in gram-negative bacteria has an impact on the fitness of the cells [30; 31]. Since TetA(B) utilizes the proton motive force to extrude the drugs, overexpression of TetA(B) in the absence of target will result in leakage of cations, which leads to loss of membrane potential and, eventually, to cell death [32; 33]. Thus, it is important that the expression levels of MDR pumps are tightly tuned to intracelllular drug concentrations, to avoid unnecessarily high, and lethal, MDR activity.

In bacteria, both transcriptional activators and repressors regulate the expression of MDR efflux pumps. Transcriptional repressors of MDR pumps include members of the TetR, MarR and LacI families, whilst transcriptional activators belong to the MerR, AraC or LysR families (Table 1). To perform their regulatory activities, many transcriptional regulators contain a DNA-binding domain and a ligand-binding domain. The ligand-binding domains of these regulators often bind the same spectrum of ligands (antibiotics or toxic compounds) as the transcriptionally controlled MDR pumps. Drug-induced upregulation of an MDR pump results from a drug-coupled conformational change in the DNA-binding domains of the transcriptional regulator, either causing a weakening (in case of a repressor) or strengthening (in case of an activator) of the DNA binding affinity. For instance, upregulation of the Bmr transporter in *Bacillus subtilis,* which is controlled by the regulator BmrR, occurs when ligand-bound BmrR binds the promoter region of the *bmr* gene. Here, BmrR acts as an activator [34]. In contrast, ligand-induced conformational changes in transcriptional repressors such as QacR in *Staphylococcus aureus* [35] and EmrR in *E. coli* [36] prevent the regulators to bind to their target promoters, leading to the upregulation of the upstream-located MDR transporter genes. All the MDR transcriptional regulators described so far (including those in Table I) are gene-specific and commonly referred as "local" transcriptional regulators.

1

Structural variations in MDR transcriptional regulators

Based on similarities of their DNA binding domains, transcriptional regulators of MDR pumps have been grouped into five families: AraC, MerR, TetR, MarR and PadR-like. All the regulators from these families have a helix-turn-helix (HTH) structural motif to bind to DNA. The two motifs in the AraC family are present in one single polypeptide chain [51]. In contrast, the two HTH motifs in the TetR family, are obtained through dimerization of two protomers harboring a single HTH motif each [52]. In MarR, MerR and PadR-like family members the DNA binding domain contains an extension to the HTH motif in the form of a "β-wing" (together referred to as a winged-HTH motif). In MerR, the wing is formed by three antiparallel β-strands, whereas the wing in the MarR and PadR family members consists of a two-stranded β-sheet [53]. In all these winged-HTH motif containing proteins, the wing is located at the C-terminus of the DNA binding domain and positioned in such a way as to provide additional interactions with the DNA [53].

In contrast to their DNA binding domains, considerable differences are observed for the ligand-binding domains of MDR-related transcriptional regulators. This is somewhat surprising, considering that many regulators bind chemically similar drugs, but highlights the importance of studying the principles of multidrug specificity. Furthermore, MDR transcriptional regulators are more feasible for structural and drug binding studies compared to membrane-bound cognate MDR pumps as they can be easily expressed and purified in high amounts and are readily crystallizable [16; 54; 55]

Table 1: Crystal structures of multidrug binding gene regulators in their drug-bound and apo-forms.

Gene regulator	Regulatory family	Target MDR pump family	Regulatory function	Ligands	PDB ID	Refs.
BmrR	MerR	MFS	Activator	Tpp, R6G, 4AQ, Ber, Kan, Ach Tet, Pur, Et	1EXI, 3Q2Y, 3Q3D, 3Q5P, 3Q5S, 3Q5R, 3Q1M, 3D6Y, 3D6Z	37, 38, 39
MexR	MarR	RND	Repressor	ArmR (antirepressor)	3ECH	40
MTH313	MarR	RND	Repressor	Salicylate	3BPV	41
MarR	MarR	RND	Repressor	Salicylate	1JGS	42
BldR	MarR	RND	Repressor	Apo-protein	3F3X	43
QacR	TetR	MFS	Repressor	Apo-protein R6G, Ber, Did Et,Prl, et, Mg Deq, Cv, Pnt		44, 45
CgmR	TetR	MFS	Repressor	Apo-protein Methylene blue Ethidium	2ZOZ, 2YVE, 2ZOY	46
TtgR	TetR	MFS	Repressor	Naringenin Quercetin Phloretin chloramphenicol Tetracycline	2UXO, 2UXP, 2UXU, 2UXH, 2UXI	47
LmrR	PadR	ABC	Repressor	Apo-protein H33342 Dm	3F8B, 3F8C, 3F8F	48
ActR	TetR	unknown	Activator	Actinorhodin, (S)-DNPA	3B6A, 3B6C	49
LfrR	TetR	MFS	Repressor	Proflavin	2WGB, 2V57	50

1) Drug abbreviations: Tetraphenylphosphonium (Tpp), Rhodamine 6G (R6G), 4-aminoquinaldine (4AQ), Berberine (Ber), Kanamycin (Kan), Acetylcholine (Ach), Tetracycline (Tet), Puromycin (Pur), Ethidium (Et), Proflavin (Prl), Crystal violet (Cv), Pentamidine (Pnt), diamidine hexamidine (Did), Hoechst33342 (H33342), Daunomycin(Dm)

As this thesis is concerned with the multidrug binding properties of the transcriptional regulator LmrR from the PadR-like family, understanding these properties for the well studied transcriptional regulators like QacR and BmrR, which belong to different families, helps in appreciating the uniqueness of the binding pockets that confer multidrug specificity. Therefore, in the following sections, structural views are presented on the multidrug binding properties of these two transcriptional regulators.

The canonical view of multidrug recognition: the multifaceted ligand-binding site of QacR of the TetR family

The *Staphylococcus aureus* MDR pump QacA, which was the first MDR pump protein reported in bacteria[56], confers resistance to monovalent and bivalent cationic lipophilic disinfectants and antiseptics such as quaternary ammonium compounds (QACs). The transcriptional regulator QacR regulates the expression of the *qacA* gene[56]. QacR was classified in the TetR family on the basis of the similarity of its N-terminal HTH motif to that of TetR[57]. The ligand/drug binding region is located at the junction between the dimerization interface of two individual monomers and the DNA binding domain (Figure 1A). QacR has a broad substrate specificity, which ranges from various quaternary ammonium compounds to disinfectants and antibiotics that represent almost all the substrates exported by the QacA pump. Several structures of QacR have been obtained, bound with different ligands, such as ethidium, rhodamine 6G, berberine, dequalinium, crystal violet and proflavin [45; 58; 59]. These ligands, which have a volume of ~300-400 $Å^3$, bind in a large voluminous drug-binding pocket of ~1100 $Å^3$. Their binding is not restricted to a common location in the drug-binding pocket, but is rather segregated into small mini-pockets (Figure 1B). Because of these distinct overlapping mini-pockets, the drug-binding pocket in QacR provides extensive and different interaction surfaces for chemically different ligands (Figure 1B).

The drug-protein interactions are dominated by interactions of the drugs with the side chains of aromatic residues (Phe, Tyr, and Trp). Aromatic residues contribute to multi-specific (alternatively referred to as promiscuous) drug binding due to their non-specific π-interactions, such as π-π, π-hydrophobic, and π-cationic, with the drugs. The binding of structurally different ligands is achieved by the different combinations of aromatic side chains and by the

flexibility of these side chains, allowing them to adopt ligand-dependent rotameric states. Due to their distribution in the voluminous drug-binding pocket and as well as due to their non-specific interactions with the planar ring system of the ligands, aromatic residues also facilitate binding of more than one drug simultaneously[44]. In addition to the aromatic side chains, H-bonding and electrostatic interactions also contribute to the plasticity in MD recognition. In QacR, at least three Glu and two Asp residues adopt drug-dependent rotameric states for optimal H-bonding interactions.

In the TtgR regulator, which belongs to same family as QacR, the binding pocket (*ca*. 1500 Å^3) is larger than that of QacR; this pocket facilitates the binding of several different, mostly plant-derived drugs with the help of hydrophobic residues [47]. However, in sharp contrast to QacR, the ligand-binding pocket in TtgR is dominated by aliphatic hydrophobic residues and contains only a single aromatic residue (six in QacR). Aliphatic hydrophobic residues are also well suited for MD recognition and are well known for their involvement

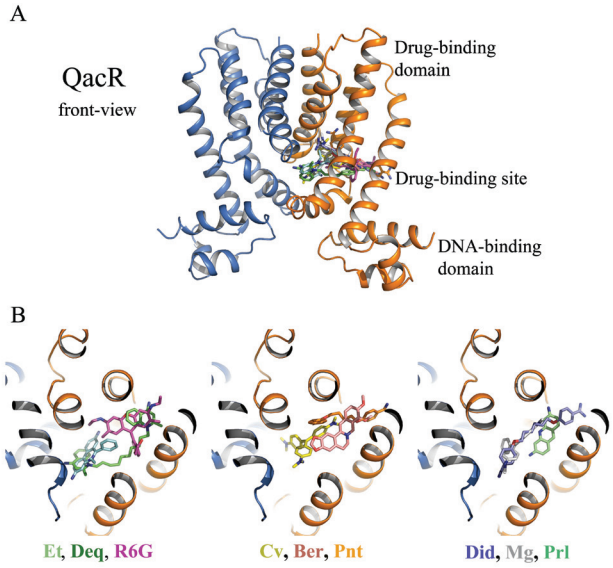

Figure 1. Canonical view of MD recognition in QacR. A) Cartoon representation of the QacR dimer from *S. aureus* (PDB entry 1JTY). The two subunits are represented in blue and orange. B) top-view of the drug binding site showing various drugs which interact with different areas of the large drug-binding cavity. Chemically different drugs are represented in sticks. Ethidium bromide (Et), dequalinium (Deq), rhodamine 6G (R6G), crystal violet (Cv), berberine (Ber), pentamidine (Pnt), diamidine hexamidine (Did), malachite green (Mg), proflavin (Prl)

in multi-specific ligand recognition like in antigen-antibody interactions[60; 61], odorant binding protein[62] and the PXR transcriptional regulator[63].

Several of the drug complexes of QacR and TtgR underscore the importance of an adaptable voluminous binding pocket and the hydrophobic character in MD recognition. These features are also shared in MD related pumps like AcrB [64] and Pgp [23] and form the basis of the canonical MD recognition features. In several other cases, like LmrA from *Lactococcus lactis* and MsbA from *E. coli,* ligand-protein interaction studies show the same canonical features [65]. Promiscuous drug recognition is not only unique to MD related proteins, but is also utilized by enzymes involved in drug metabolism and disposition[66]. For example, in the drug-metabolizing enzyme cytochrome P450 the ligand-binding site is similar to that of QacR: it is a voluminous ligand-binding site with aromatic residues that can bind more than one drug simultaneously [67].

A non-canonical view of multi-drug recognition in BmrR of the MerR family

The transcriptional activator BmrR regulates the expression of the MDR transporter Bmr in *Bacillus subtilis* [34]. Like QacR, BmrR is functionally active as a homodimer. Within each monomer, the ligand binding domain and the DNA binding domain are far away from each other (Figure 2A). However, in the dimeric form, the ligand binding domain of one monomer interacts with the DNA binding domain of the other, facilitating the drug induced activation of the *bmr* promoter [39]. Because of this structural organization, there are two identical drug-binding sites in the BmrR dimer (Figure 2A). The drug-binding site of BmrR (*ca.* 650 Å^3) is much smaller than that of QacR (*ca*. 1100Å^3). It lacks distinct multiple sites and an adaptable binding pocket (Figure 2B), which are the core requirements for the canonical drug binding site as described above. In BmrR, the drug-binding site is rather rigid and offers a platform for chemically unrelated ligands by retaining key MD-binding elements: aromatic and aliphatic residues. In several of the BmrR drug-bound complexes [37], these residues form a conformationally rigid ring-like scaffold for binding the aromatic ring systems common to the various ligands. Another subsite, which is more solvent exposed and contains hydrogen-bonding elements, provides a hydrophilic cavity for binding ligands lacking rigid or planar moieties.

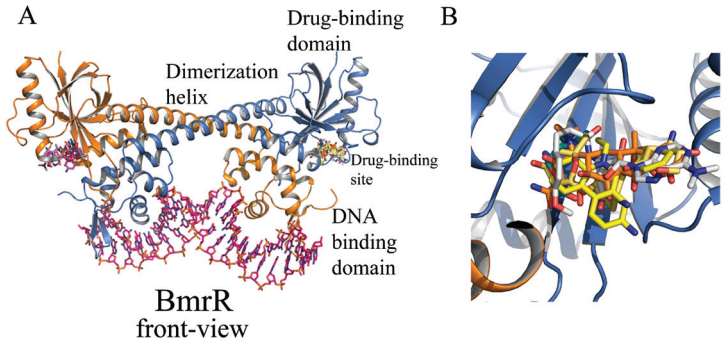

Figure 2. BmrR model and non-canonical MD recognition. A) Cartoon representation of the structure of BmrR from *B. subtilis* in the presence of DNA (PDB entry 1EXJ). The two equivalent ligand-binding sites (dyad-related) contain an overlay of several ligands (shown as sticks in yellow and pink). B) Close-up side view of the drug-binding site. Various drugs bound at a common surface are depicted as sticks.

Transcriptional regulators of the PadR family

Microorganisms living in a plant-soil ecosystem survive phenolic acid stress by the expression of a phenolic acid decarboxylase gene (*padA* gene). Phenolic acids are abundant in plant cell walls and vacuoles [68]. They are released as free acids such as *p*-coumaric, ferulic and caffeic acids by the action of hemicellulases produced by fungi and bacteria. These antimicrobial compounds are converted into less-toxic derivatives by phenolic acid decarboxylase (Pad) encoded by the inducible *padA* gene[69; 70]. The transcriptional regulator PadR negatively regulates the expression of the *padA* gene. PadR was the founding member of a new class of transcriptional regulators, the PadR family (Pfam PF03551). The PadR family now includes a large and diverse group of proteins, of which only a few have been characterized. Among these proteins, the transcriptional activator AphA (31% identity to PadR) initiates the virulence cascade in *Vibrio cholerae* [71; 72], without apparently playing a role in the regulation of Pad [73]. LadR is a regulator that is widely conserved in firmicutes and related organisms; it regulates the expression of the multidrug efflux pump MdrL [74]. This efflux pump is known to extrude multiple substrates, which include ethidium, macrolides, cefotaxime and heavy metals [75]. Other members of the PadR family are LstR, which is required for effective thermal resistance [76], and LmrR from *Lactococcus lactis,* which regulates the production of LmrCD, a major multidrug

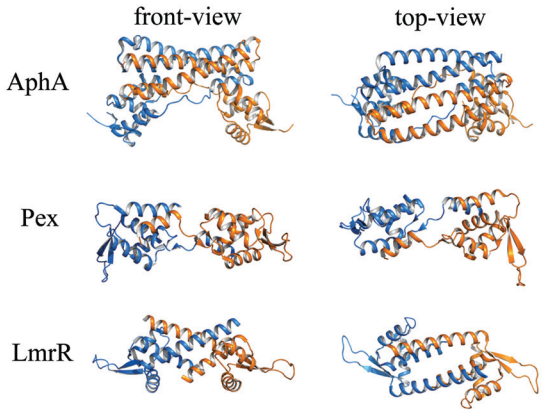

Figure 3. Variations in the dimerization domains of different PadR family members. Cartoon representation of AphA (PDB entry 1YG2), Pex (PDB entry 2E1N) and LmrR (PDB entry 3F8B). The two subunits of each functional form are represented in two different colors (blue and orange).

ABC transporter [48; 77]. Based on the length of their ORFs, *padR*-related genes have been classified into two distinct subfamilies, subfamilies I and II, which encode proteins of ~176 and ~110 amino acids, respectively[74]. PadR, LadR and AphA belong to the PadR subfamily I and LmrR belongs to subfamily II. Crystal structures of two PadR-like proteins, AphA [78] and Pex [79], revealed the presence of an N-terminal wHTH DNA binding motif, which is highly similar to that found in the MarR family[42]. However, the C-terminal dimerization domains of AphA and Pex are highly divergent in PadR as well as from the MarR family (top-view in Figure 3).

LmrR from *Lactococcus lactis*

The *lmrR* (*l*actococcal *m*ultidrug *r*esistance *r*egulator) gene is located upstream of the *lmrCD* genes. LmrR, the expression product of the *lmrR* gene, negatively regulates the expression of the ABC transporter LmrCD, a major determinant of the MDR phenotype of *Lactococcus lactis* [80]. In non-induced conditions, LmrR binds to the promoter of *lmrCD,* thereby repressing its expression. In the presence of inducers, like the cellular toxic compounds daunomycin, ethidium, Hoechst 33342, and the fluorescence dye BCECF-AM (5-carboxyfluorescein, diacetoxymethyl ester), repression of *lmrCD* is relieved resulting in increased

production of the membrane-bound efflux pump LmrCD. Possibly, LmrR undergoes a conformational change upon drug binding, which prohibits it from binding to the promoter of the *lmrCD* genes. At the same time, in *L. lactis* MDR strains, it has been observed that expression of LmrR itself is also enhanced[81]. Thus LmrR plays a dual regulatory role by binding, though in a different stoichiometry, to the control regions of both the *lmrR* and *lmrCD* genes[82]. LmrR is bound to its own control region in a high copy number fashion, whereas two copies of LmrR dimer are bound to the *lmrCD* control region. This might ensure a demand-based expression of LmrCD. For example, at elevated levels of toxicity both the regulator and the functional genes could be induced and co-transcribed from a polycistronic messenger, and at the same time a single transcript of LmrCD could be transcribed leading to the higher-level expression of LmrCD.

Crystal structures of apo LmrR, and complexes of LmrR with bound Hoechst 33342 and daunomycin have been solved at 2.0 Å and 2.2 Å, respectively [48]. Remarkably, although LmrR retains an N-terminal helix-turn-helix motif for binding to DNA, its C-terminal ligand(s) binding domain has a unique architecture that deviates from the well studied QacR and BmrR regulators. The dimer contains a large central pore at the subunit interface that is symmetric with equal contributions from both monomers. The planar ring system of the two toxic compounds is wedged in between the W96 and W96' side chains forming aromatic stacking interactions with each of the two indole systems. There are no hydrogen bonding interactions between the drug(s) and LmrR.

Concluding Remarks

Pathogenic microorganisms have evolved over time to survive the continuous onslaught of antibiotics. Due to the abuse of antibiotics, several microbes, which once were sensitive, are transforming into 'superbugs' by becoming resistant to almost all available antibiotics. Resistance comes in different flavors and is problematic to investigate. This is caused, firstly, by the different types of resistance mechanisms, secondly because pathogenic microorganisms employ a combination of resistance mechanisms to ensure their safety from a plethora of antibiotics, and, finally, because these mechanisms can be transferred rapidly to non-resistant microorganisms. One way to accelerate

the production of new antibiotics is to understand the mechanisms by which pathogenic microbes attain resistance. For such an approach multidrug efflux pumps are of significant importance because of their ability to recognize the chemically different toxic compounds by which microbes attain multidrug resistance. However, structural studies on efflux pumps are hampered by the difficulties in purifying and crystallizing membrane proteins. On the other hand, MDR related regulatory proteins are soluble, relatively easy to produce in the amounts required for structural studies and it is often easy to obtain diffracting crystals. Structural studies on regulatory proteins are beginning to provide features for multidrug recognition. For example, numerous drug-bound complexes of QacR and TtgR display the canonical basis of multi-drug recognition, which is also shared by MDR-related pumps like Pgp and AcrB. In addition to this model, BmrR drug bound complexes presented a non-canonical model of drug recognition. These studies now indicate that multidrug recognition can be varied. Our current knowledge about different MD recognition mechanisms is limited. There is a further need for additional structural solutions from different families to enrich the understanding of MD recognition.

References

1. Forrest, R. D. (1982). Early history of wound treatment. *J R Soc Med* **75**, 198-205.
2. Taubes, G. (2008). The bacteria fight back. *Science* **321**, 356-61.
3. Levy, S. B. & Marshall, B. (2004). Antibacterial resistance worldwide: causes, challenges and responses. *Nat Med* **10**, S122-9.
4. Wang, H., Dzink-Fox, J. L., Chen, M. & Levy, S. B. (2001). Genetic characterization of highly fluoroquinolone-resistant clinical Escherichia coli strains from China: role of acrR mutations. *Antimicrob Agents Chemother* **45**, 1515-21.
5. Schneiders, T., Amyes, S. G. & Levy, S. B. (2003). Role of AcrR and ramA in fluoroquinolone resistance in clinical Klebsiella pneumoniae isolates from Singapore. *Antimicrob Agents Chemother* **47**, 2831-7.
6. Barza, M., Gorbach, S. & DeVincent, S. J. (2002). The need to improve antimicrobial use in agriculture: ecological and human health consequences. *Clinical Infectious Diseases* **34**, S71-S144.
7. Carlet, J., Collignon, P., Goldmann, D., Goossens, H., Gyssens, I. C., Harbarth, S., Jarlier, V., Levy, S. B., N'Doye, B., Pittet, D., Richtmann, R., Seto, W. H., van der Meer, J. W. & Voss, A. (2011). Society's failure to protect a precious resource: antibiotics. *Lancet* **378**, 369-71.
8. Levy, S. B. (1998). The challenge of antibiotic resistance. *Sci Am* **278**, 46-53.
9. Walsh, F. M. & Amyes, S. G. (2004). Microbiology and drug resistance mechanisms of fully resistant pathogens. *Curr Opin Microbiol* **7**, 439-44.
10. Iseman, M. D. (1993). Treatment of multidrug-resistant tuberculosis. *N Engl J Med* **329**, 784-91.
11. Hegreness, M., Shoresh, N., Damian, D., Hartl, D. & Kishony, R. (2008). Accelerated evolution of resistance in multidrug environments. *Proc Natl Acad Sci U S A* **105**, 13977-81.
12. Wright, G. D. (1999). Aminoglycoside-modifying enzymes. *Curr Opin Microbiol* **2**, 499-503.
13. Connell, S. R., Tracz, D. M., Nierhaus, K. H. & Taylor, D. E. (2003). Ribosomal protection proteins and their mechanism of tetracycline resistance. *Antimicrob Agents Chemother* **47**, 3675-81.
14. Kohanski, M. A., DePristo, M. A. & Collins, J. J. (2010). Sublethal antibiotic treatment leads to multidrug resistance via radical-induced mutagenesis. *Mol Cell* **37**, 311-20.
15. Beaber, J. W., Hochhut, B. & Waldor, M. K. (2004). SOS response promotes horizontal dissemination of antibiotic resistance genes. *Nature* **427**, 72-4.
16. Alekshun, M. N. & Levy, S. B. (2007). Molecular mechanisms of antibacterial multidrug resistance. *Cell* **128**, 1037-50.
17. Poole, K. (2004). Efflux-mediated multiresistance in Gram-negative bacteria. *Clin Microbiol Infect* **10**, 12-26.
18. Paulsen, I. T., Brown, M. H. & Skurray, R. A. (1998). Characterization of the earliest known Staphylococcus aureus plasmid encoding a multidrug efflux system. *J Bacteriol* **180**, 3477-9.
19. Paulsen, I. T., Brown, M. H. & Skurray, R. A. (1996). Proton-dependent multidrug efflux systems. *Microbiol Rev* **60**, 575-608.

20. Lubelski, J., Konings, W. N. & Driessen, A. J. (2007). Distribution and physiology of ABC-type transporters contributing to multidrug resistance in bacteria. *Microbiol Mol Biol Rev* **71**, 463-76.
21. Jack, D. L., Yang, N. M. & Saier, M. H., Jr. (2001). The drug/metabolite transporter superfamily. *Eur J Biochem* **268**, 3620-39.
22. Li, X. Z. & Nikaido, H. (2004). Efflux-mediated drug resistance in bacteria. *Drugs* **64**, 159-204.
23. Aller, S. G., Yu, J., Ward, A., Weng, Y., Chittaboina, S., Zhuo, R., Harrell, P. M., Trinh, Y. T., Zhang, Q., Urbatsch, I. L. & Chang, G. (2009). Structure of P-glycoprotein reveals a molecular basis for poly-specific drug binding. *Science* **323**, 1718-22.
24. Murakami, S., Nakashima, R., Yamashita, E., Matsumoto, T. & Yamaguchi, A. (2006). Crystal structures of a multidrug transporter reveal a functionally rotating mechanism. *Nature* **443**, 173-9.
25. Ubarretxena-Belandia, I., Baldwin, J. M., Schuldiner, S. & Tate, C. G. (2003). Three-dimensional structure of the bacterial multidrug transporter EmrE shows it is an asymmetric homodimer. *EMBO J* **22**, 6175-81.
26. Ma, D., Cook, D. N., Alberti, M., Pon, N. G., Nikaido, H. & Hearst, J. E. (1995). Genes acrA and acrB encode a stress-induced efflux system of Escherichia coli. *Mol Microbiol* **16**, 45-55.
27. Ahmed, M., Lyass, L., Markham, P. N., Taylor, S. S., Vazquez-Laslop, N. & Neyfakh, A. A. (1995). Two highly similar multidrug transporters of Bacillus subtilis whose expression is differentially regulated. *J Bacteriol* **177**, 3904-10.
28. Woolridge, D. P., Vazquez-Laslop, N., Markham, P. N., Chevalier, M. S., Gerner, E. W. & Neyfakh, A. A. (1997). Efflux of the natural polyamine spermidine facilitated by the Bacillus subtilis multidrug transporter Blt. *J Biol Chem* **272**, 8864-6.
29. Levy, S. B. (1992). Active efflux mechanisms for antimicrobial resistance. *Antimicrob Agents Chemother* **36**, 695-703.
30. Nguyen, T. N., Phan, Q. G., Duong, L. P., Bertrand, K. P. & Lenski, R. E. (1989). Effects of carriage and expression of the Tn10 tetracycline-resistance operon on the fitness of Escherichia coli K12. *Mol Biol Evol* **6**, 213-25.
31. Lee, S. W. & Edlin, G. (1985). Expression of tetracycline resistance in pBR322 derivatives reduces the reproductive fitness of plasmid-containing Escherichia coli. *Gene* **39**, 173-80.
32. Eckert, B. & Beck, C. F. (1989). Overproduction of transposon Tn10-encoded tetracycline resistance protein results in cell death and loss of membrane potential. *J Bacteriol* **171**, 3557-9.
33. Hickman, R. K., McMurry, L. M. & Levy, S. B. (1990). Overproduction and purification of the Tn10-specified inner membrane tetracycline resistance protein Tet using fusions to beta-galactosidase. *Mol Microbiol* **4**, 1241-51.
34. Ahmed, M., Borsch, C. M., Taylor, S. S., Vazquez-Laslop, N. & Neyfakh, A. A. (1994). A protein that activates expression of a multidrug efflux transporter upon binding the transporter substrates. *J Biol Chem* **269**, 28506-13.
35. Grkovic, S., Brown, M. H., Roberts, N. J., Paulsen, I. T. & Skurray, R. A. (1998). QacR is a repressor protein that regulates expression of the Staphylococcus aureus multidrug efflux pump QacA. *J Biol Chem* **273**, 18665-73.

36. Lomovskaya, O., Lewis, K. & Matin, A. (1995). EmrR is a negative regulator of the Escherichia coli multidrug resistance pump EmrAB. *J Bacteriol* **177**, 2328-34.
37. Bachas, S., Eginton, C., Gunio, D. & Wade, H. (2011). Structural contributions to multidrug recognition in the multidrug resistance (MDR) gene regulator, BmrR. *Proc Natl Acad Sci U S A* **108**, 11046-51.
38. Newberry, K. J., Huffman, J. L., Miller, M. C., Vazquez-Laslop, N., Neyfakh, A. A. & Brennan, R. G. (2008). Structures of BmrR-drug complexes reveal a rigid multidrug binding pocket and transcription activation through tyrosine expulsion. *J Biol Chem* **283**, 26795-804.
39. Heldwein, E. E. & Brennan, R. G. (2001). Crystal structure of the transcription activator BmrR bound to DNA and a drug. *Nature* **409**, 378-82.
40. Wilke, M. S., Heller, M., Creagh, A. L., Haynes, C. A., McIntosh, L. P., Poole, K. & Strynadka, N. C. (2008). The crystal structure of MexR from Pseudomonas aeruginosa in complex with its antirepressor ArmR. *Proc Natl Acad Sci U S A* **105**, 14832-7.
41. Saridakis, V., Shahinas, D., Xu, X. & Christendat, D. (2008). Structural insight on the mechanism of regulation of the MarR family of proteins: high-resolution crystal structure of a transcriptional repressor from Methanobacterium thermoautotrophicum. *J Mol Biol* **377**, 655-67.
42. Alekshun, M. N., Levy, S. B., Mealy, T. R., Seaton, B. A. & Head, J. F. (2001). The crystal structure of MarR, a regulator of multiple antibiotic resistance, at 2.3 A resolution. *Nat Struct Biol* **8**, 710-4.
43. Di Fiore, A., Fiorentino, G., Vitale, R. M., Ronca, R., Amodeo, P., Pedone, C., Bartolucci, S. & De Simone, G. (2009). Structural analysis of BldR from Sulfolobus solfataricus provides insights into the molecular basis of transcriptional activation in Archaea by MarR family proteins. *J Mol Biol* **388**, 559-69.
44. Schumacher, M. A., Miller, M. C. & Brennan, R. G. (2004). Structural mechanism of the simultaneous binding of two drugs to a multidrug-binding protein. *EMBO J* **23**, 2923-30.
45. Murray, D. S., Schumacher, M. A. & Brennan, R. G. (2004). Crystal structures of QacR-diamidine complexes reveal additional multidrug-binding modes and a novel mechanism of drug charge neutralization. *J Biol Chem* **279**, 14365-71.
46. Itou, H., Watanabe, N., Yao, M., Shirakihara, Y. & Tanaka, I. Crystal structures of the multidrug binding repressor Corynebacteriumglutamicum CgmR in complex with inducers and with an operator. *J Mol Biol* **403**, 174-84.
47. Alguel, Y., Meng, C., Teran, W., Krell, T., Ramos, J. L., Gallegos, M. T. & Zhang, X. (2007). Crystal structures of multidrug binding protein TtgR in complex with antibiotics and plant antimicrobials. *J Mol Biol* **369**, 829-40.
48. Madoori, P. K., Agustiandari, H., Driessen, A. J. & Thunnissen, A. M. (2009). Structure of the transcriptional regulator LmrR and its mechanism of multidrug recognition. *EMBO J* **28**, 156-66.
49. Willems, A. R., Tahlan, K., Taguchi, T., Zhang, K., Lee, Z. Z., Ichinose, K., Junop, M. S. & Nodwell, J. R. (2008). Crystal structures of the Streptomyces coelicolor TetR-like protein ActR alone and in complex with actinorhodin or the actinorhodin biosynthetic precursor (S)-DNPA. *J Mol Biol* **376**, 1377-87.
50. Bellinzoni, M., Buroni, S., Schaeffer, F., Riccardi, G., De Rossi, E. & Alzari, P. M. (2009). Structural plasticity and distinct drug-binding modes of LfrR, a mycobacterial efflux pump regulator. *J Bacteriol* **191**, 7531-7.

51. Kwon, H. J., Bennik, M. H., Demple, B. & Ellenberger, T. (2000). Crystal structure of the Escherichia coli Rob transcription factor in complex with DNA. *Nat Struct Biol* **7**, 424-30.
52. Orth, P., Schnappinger, D., Hillen, W., Saenger, W. & Hinrichs, W. (2000). Structural basis of gene regulation by the tetracycline inducible Tet repressor-operator system. *Nat Struct Biol* **7**, 215-9.
53. Aravind, L., Anantharaman, V., Balaji, S., Babu, M. M. & Iyer, L. M. (2005). The many faces of the helix-turn-helix domain: transcription regulation and beyond. *FEMS Microbiol Rev* **29**, 231-62.
54. Nikaido, H. (2009). Multidrug resistance in bacteria. *Annu Rev Biochem* **78**, 119-46.
55. Higgins, C. F. (2007). Multiple molecular mechanisms for multidrug resistance transporters. *Nature* **446**, 749-57.
56. Tennent, J. M., Lyon, B. R., Gillespie, M. T., May, J. W. & Skurray, R. A. (1985). Cloning and expression of Staphylococcus aureus plasmid-mediated quaternary ammonium resistance in Escherichia coli. *Antimicrob Agents Chemother* **27**, 79-83.
57. Rouch, D. A., Cram, D. S., DiBerardino, D., Littlejohn, T. G. & Skurray, R. A. (1990). Efflux-mediated antiseptic resistance gene qacA from Staphylococcus aureus: common ancestry with tetracycline- and sugar-transport proteins. *Mol Microbiol* **4**, 2051-62.
58. Schumacher, M. A., Miller, M. C., Grkovic, S., Brown, M. H., Skurray, R. A. & Brennan, R. G. (2001). Structural mechanisms of QacR induction and multidrug recognition. *Science* **294**, 2158-63.
59. Brooks, B. E., Piro, K. M. & Brennan, R. G. (2007). Multidrug-binding transcription factor QacR binds the bivalent aromatic diamidines DB75 and DB359 in multiple positions. *J Am Chem Soc* **129**, 8389-95.
60. Padlan, E. A. (1990). On the nature of antibody combining sites: unusual structural features that may confer on these sites an enhanced capacity for binding ligands. *Proteins* **7**, 112-24.
61. Mariuzza, R. A., Phillips, S. E. & Poljak, R. J. (1987). The structural basis of antigen-antibody recognition. *Annu Rev Biophys Biophys Chem* **16**, 139-59.
62. Bianchet, M. A., Bains, G., Pelosi, P., Pevsner, J., Snyder, S. H., Monaco, H. L. & Amzel, L. M. (1996). The three-dimensional structure of bovine odorant binding protein and its mechanism of odor recognition. *Nat Struct Biol* **3**, 934-9.
63. Watkins, R. E., Wisely, G. B., Moore, L. B., Collins, J. L., Lambert, M. H., Williams, S. P., Willson, T. M., Kliewer, S. A. & Redinbo, M. R. (2001). The human nuclear xenobiotic receptor PXR: structural determinants of directed promiscuity. *Science* **292**, 2329-33.
64. Yu, E. W., McDermott, G., Zgurskaya, H. I., Nikaido, H. & Koshland, D. E., Jr. (2003). Structural basis of multiple drug-binding capacity of the AcrB multidrug efflux pump. *Science* **300**, 976-80.
65. Gutmann, D. A., Ward, A., Urbatsch, I. L., Chang, G. & van Veen, H. W. (2010). Understanding polyspecificity of multidrug ABC transporters: closing in on the gaps in ABCB1. *Trends Biochem Sci* **35**, 36-42.
66. Ma, Q. & Lu, A. Y. (2008). The challenges of dealing with promiscuous drug-metabolizing enzymes, receptors and transporters. *Curr Drug Metab* **9**, 374-83.
67. Cupp-Vickery, J., Anderson, R. & Hatziris, Z. (2000). Crystal structures of ligand complexes of P450eryF exhibiting homotropic cooperativity. *Proc Natl Acad Sci U S A* **97**, 3050-5.

68. Lynd, L. R., Weimer, P. J., van Zyl, W. H. & Pretorius, I. S. (2002). Microbial cellulose utilization: fundamentals and biotechnology. *Microbiol Mol Biol Rev* **66**, 506-77, table of contents.
69. Barthelmebs, L., Divies, C. & Cavin, J. F. (2000). Knockout of the p-coumarate decarboxylase gene from Lactobacillus plantarum reveals the existence of two other inducible enzymatic activities involved in phenolic acid metabolism. *Appl Environ Microbiol* **66**, 3368-75.
70. Cavin, J. F., Barthelmebs, L. & Divies, C. (1997). Molecular characterization of an inducible p-coumaric acid decarboxylase from Lactobacillus plantarum: gene cloning, transcriptional analysis, overexpression in Escherichia coli, purification, and characterization. *Appl Environ Microbiol* **63**, 1939-44.
71. Kovacikova, G. & Skorupski, K. (2001). Overlapping binding sites for the virulence gene regulators AphA, AphB and cAMP-CRP at the Vibrio cholerae tcpPH promoter. *Mol Microbiol* **41**, 393-407.
72. Kovacikova, G. & Skorupski, K. (1999). A Vibrio cholerae LysR homolog, AphB, cooperates with AphA at the tcpPH promoter to activate expression of the ToxR virulence cascade. *J Bacteriol* **181**, 4250-6.
73. Kovacikova, G., Lin, W. & Skorupski, K. (2003). The virulence activator AphA links quorum sensing to pathogenesis and physiology in Vibrio cholerae by repressing the expression of a penicillin amidase gene on the small chromosome. *J Bacteriol* **185**, 4825-36.
74. Huillet, E., Velge, P., Vallaeys, T. & Pardon, P. (2006). LadR, a new PadR-related transcriptional regulator from Listeria monocytogenes, negatively regulates the expression of the multidrug efflux pump MdrL. *FEMS Microbiol Lett* **254**, 87-94.
75. Mata, M. T., Baquero, F. & Perez-Diaz, J. C. (2000). A multidrug efflux transporter in Listeria monocytogenes. *FEMS Microbiol Lett* **187**, 185-8.
76. Zhang, C., Nietfeldt, J., Zhang, M. & Benson, A. K. (2005). Functional consequences of genome evolution in Listeria monocytogenes: the lmo0423 and lmo0422 genes encode sigmaC and LstR, a lineage II-specific heat shock system. *J Bacteriol* **187**, 7243-53.
77. Agustiandari, H., Lubelski, J., van den Berg van Saparoea, H. B., Kuipers, O. P. & Driessen, A. J. (2008). LmrR is a transcriptional repressor of expression of the multidrug ABC transporter LmrCD in Lactococcus lactis. *J Bacteriol* **190**, 759-63.
78. De Silva, R. S., Kovacikova, G., Lin, W., Taylor, R. K., Skorupski, K. & Kull, F. J. (2005). Crystal structure of the virulence gene activator AphA from Vibrio cholerae reveals it is a novel member of the winged helix transcription factor superfamily. *J Biol Chem* **280**, 13779-83.
79. Arita, K., Hashimoto, H., Igari, K., Akaboshi, M., Kutsuna, S., Sato, M. & Shimizu, T. (2007). Structural and biochemical characterization of a cyanobacterium circadian clock-modifier protein. *J Biol Chem* **282**, 1128-35.
80. Lubelski, J., Mazurkiewicz, P., van Merkerk, R., Konings, W. N. & Driessen, A. J. (2004). ydaG and ydbA of Lactococcus lactis encode a heterodimeric ATP-binding cassette-type multidrug transporter. *J Biol Chem* **279**, 34449-55.
81. Lubelski, J., de Jong, A., van Merkerk, R., Agustiandari, H., Kuipers, O. P., Kok, J. & Driessen, A. J. (2006). LmrCD is a major multidrug resistance transporter in Lactococcus lactis. *Mol Microbiol* **61**, 771-81.

82. Agustiandari, H., Peeters, E., de Wit, J. G., Charlier, D. & Driessen, A. J. (2011). LmrR-mediated gene regulation of multidrug resistance in Lactococcus lactis. *Microbiology* **157**, 1519-30.

Chapter

Structure of the transcriptional regulator LmrR and its mechanism of multi-drug recognition

2

Pramod Kumar Madoori, Herfita Agustiandari, Arnold J. M. Driessen and Andy-Mark W. H. Thunnissen

Published in: *EMBO Journal* (2008) **28**, 156 - 166

Abstract

LmrR is a PadR-related transcriptional repressor that regulates the production of LmrCD, a major multidrug ABC transporter in *Lactococcus lactis*. Transcriptional regulation is presumed to follow a drug-sensitive induction mechanism involving the direct binding of transporter ligands to LmrR. Here we present crystal structures of LmrR in the apo state and in two drug-bound states, complexed with Hoechst 33342 and daunomycin, respectively. LmrR shows a common topology containing a typical β-winged helix-turn-helix domain with an additional C-terminal helix involved in dimerization. Its dimeric organization is highly unusual with a flat-shaped hydrophobic pore at the dimer centre serving as a multi-drug binding site. The drugs bind in a similar fashion with their aromatic rings sandwiched in between the indole groups of two dimer-related tryptophan residues. Multi-drug recognition is facilitated by conformational plasticity and the absence of drug-specific hydrogen bonds. Combined analyses using site-directed mutagenesis, fluorescence-based drug binding and protein-DNA gel shift assays reveal an allosteric coupling between the multidrug and DNA binding sites of LmrR that likely plays a role in the induction mechanism.

Introduction

Multidrug resistance (MDR) is frequently caused by the action of specialized membrane-bound pumps that possess or have acquired the ability to extrude a wide variety of chemically and structurally different compounds from the cell (Higgins, 2007; Saier et al, 1998). The molecular mechanisms of substrate recognition by these multi-drug transporters are poorly understood, mainly because the proteins involved are very recalcitrant towards crystallization, a prerequisite for a detailed structural analysis by X-ray crystallography. Instead, general features explaining multi-drug binding specificity have been derived from structural studies of the transcriptional regulators of multi-drug transporters, which are soluble proteins and often bind many of the same drugs that are substrates of the pumps (Alguel, 2007; Heldwein & Brennan, 2001; Higgins, 2007; Schumacher et al, 2001; Zheleznova et al, 1999).

LmrR is a recently identified transcription factor that controls the expression of the heterodimeric ABC transporter LmrCD, which is a major multidrug transporter in *Lactococcus lactis* (Lubelski et al, 2006). It is encoded in the immediate vicinity of the *lmrCD* genes and was shown to specifically bind to the *lmrCD* and *lmrR* promoters where it acts as a transcriptional repressor and autoregulator, respectively (Agustiandari et al, 2008). Toxic compounds that form substrates of the LmrCD transporter include the DNA-binding drugs Hoechst 33342, daunomycin, ethidium bromide and rhodamine 6G. One of these drugs, Hoechst 33342, was shown to directly interact with LmrR, and its presence in the growth medium induced a significant up-regulation of the *lmrCD* genes (Agustiandari et al, 2008). This strongly suggests that the transcription factor may act as a drug-sensor causing stimulation of LmrCD production in the presence of toxic compounds, thus promoting their extrusion from the cell. By homology LmrR belongs to the PadR family of transcriptional regulators found in bacteria and archaea (Gury et al, 2004; Huillet et al, 2006). Only a few PadR family members have been functionally characterized, showing that these proteins play important roles in the regulation of distinct cellular pathways leading, for example, to MDR, virulence and detoxification. Crystal structures of two PadR-like proteins, AphA (De Silva et al, 2005) and Pex (Arita et al, 2007), revealed the characteristics of the PadR fold, which includes a conserved N-terminal winged helix-turn-helix (wHTH) DNA binding domain (Aravind et al, 2005) that is architecturally similar to that of the multiple

antibiotic resistance repressor MarR (Alekshun et al, 2001), and a highly diverse C-terminal helical domain which serves as a dimerization module.

No structural data is available on multi-drug binding transcriptional regulators of the PadR family. To understand the structural basis for multi-drug recognition by LmrR, as well as to provide insights into the molecular mechanisms involved in the regulation of *lmrCD* expression, we have determined the crystal structure of LmrR in the apo state, as well as in two drug-bound states, complexed with Hoechst 33342 and daunomycin, respectively.

RESULTS

Structure Determination of LmrR in Apo and Drug-Bound States

Full length LmrR was purified both as a fusion protein containing a C-terminal streptactin-tag and as untagged protein. The crystal structure of untagged LmrR was determined to 2.0 Å resolution, by the molecular replacement method using an ensemble of three structurally homologous, but functionally uncharacterized proteins from the Protein Data Bank (PDB) as a search model. The untagged LmrR structure was solved in the absence of bound drugs and thus represents the apo form. Tagged LmrR was also crystallized in the absence of drugs, and its structure was determined at 2.5 Å resolution. Since the overall structural features of drug-free tagged LmrR are identical to those of drug-free untagged LmrR, we will describe the latter structure only and refer to it as apo-LmrR.

Crystals of drug-bound LmrR were obtained by co-crystallization using tagged and untagged LmrR preincubated with different lipophilic cationic drugs. Co-crystals of tagged LmrR complexed with Hoechst 33342 (H33342) and with daunomycin diffracted both to 2.2 Å resolution. Co-crystals obtained with untagged LmrR, or in the presence of ethidium bromide and rhodamine 6G, could not be used for structure determination because of their poor X-ray diffraction quality. In all cases, however, the drug was present as indicated by a drug-specific coloring of the crystals. The structures of LmrR bound to H33342 and daunomycin were determined by molecular replacement using the apo-LmrR subunit structure as a search model.

Each of the three LmrR structures (apo, H33342-bound and daunomycin-bound) represents a different crystal form with one or two LmrR subunits in the asymmetric unit. In all crystals LmrR is present as dimers formed via either crystallographic or non-crystallographic symmetry. The dimeric nature of LmrR is consistent with the results of gel filtration chromatography and dynamic light scattering experiments (not shown), and agrees with the general oligomeric preference of other MarR/PadR family members. The overall fold of the LmrR dimer is the same in the different structures, but structural superpositions reveal some notable differences, highlighting a significant inherent conformational plasticity (further explained below). The stereochemical quality of the models is excellent with no Ramachandran outliers. The electron density for the polypeptide chains is generally well defined, except for the N- and C-termini, including the strep-tag, which are disordered and not included in the final models. Another flexible region that is disordered in most of the LmrR structures comprises the β-wing loop (residues 70-75). The drug-binding site is located at the dyad axis of the LmrR dimer. In the LmrR-H33342 complex structure, where the dimer is formed by crystallographic symmetry, the Hoechst compound binds in two mutually exclusive orientations related by the crystallographic dyad symmetry, resulting in an averaged electron density at the drug binding site (Supplementary Figure S1A). However, electron density for H33342 is well defined, allowing straightforward deconvolution of the two symmetry-related binding modes. Electron density for the daunomycin molecule in the LmrR-daunomycin complex is less well defined (Supplementary Figure S1B), indicating some disorder in binding. A summary of the data collection and model refinement statistics is presented in Table 1.

Overall Structure of Apo-LmrR

The apo-LmrR crystal contains two independent copies of a subunit in the asymmetric unit, which, through a crystallographic dyad rotation axis, form a biologically relevant dimer with approximate overall dimensions of 100 Å × 38 Å × 38 Å (Figure 1). Each LmrR subunit has an (α + β) structure with topology α1 (residues 6-23), α2 (residues 28-39), α3 (residues 47-60), β1 (residues 63-67), β2 (residues 77-81), α4 (residues 83-108), and is divided into two functional domains: a typical wHTH DNA binding domain, which consists of helices α1, α2, the DNA recognition helix α3 and strands β1 and β2 (together forming the

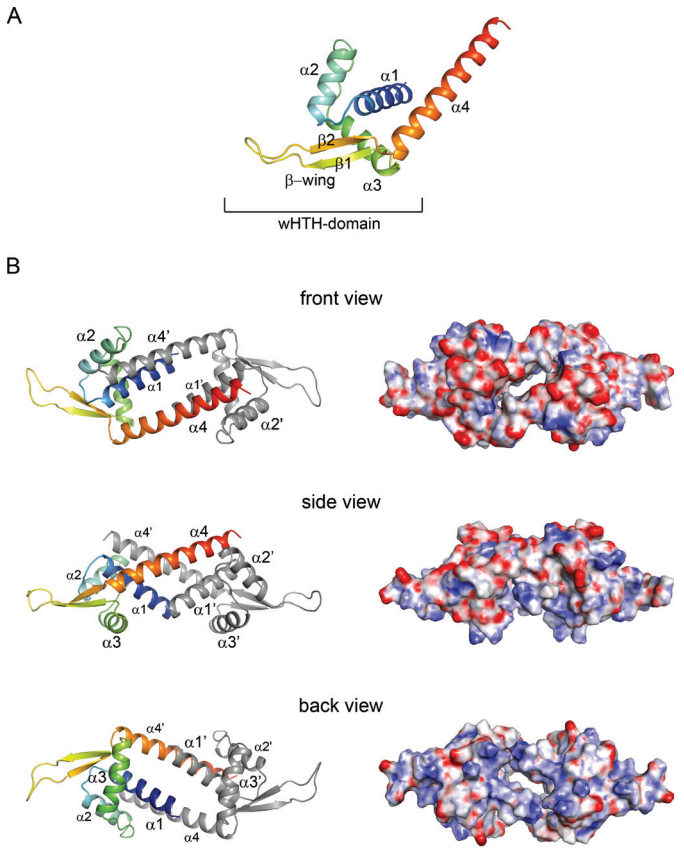

Figure 1. Overall structure of apo-LmrR (A) Ribbon representation of a single LmrR subunit with a rainbow color gradient from the N-terminus (blue) to the C-terminus (red). Secondary structure elements are indicated with labels. (B) The apo-LmrR dimer is shown in three orientations, related by 90° rotations, resulting in a front view (along the 2-fold rotation axis facing the α4 helices), a side view (perpendicular to the 2-fold axis) and a back view (along the 2-fold axis facing the α1 and α3 helices). Helices are indicated with labels. The left panel shows the LmrR dimer in a ribbon presentation, the right panel in electrostatic surface representation. The red and blue in the surface representations indicate strength of electrostatic surface potential (red, negative charge; blue, positive charge).

wing), and a dimerization domain containing the C-terminal helix α4 (Figure 1A). Helix α4 forms a protruding arm, which in the dimer crosses over to the wHTH domain of the dyad-related subunit, packing in a nearly antiparallel orientation against helix α1' (the prime indicates the other subunit), as well

as interacting with the C-terminal region of helix α2' and the loop connecting helices α2' and α3' (Figure 1B). Although facing each other at the centre of the dimer, there is no interaction between the C-terminal helices α4 and α4', nor between N-terminal helices α1 and α1'. Remarkably, this dimeric arrangement results in the formation of a large flat-shaped pore (approximately 22 Å in width and 6 Å in height) running through the dimer and centered around the dyad. The pore entrances are formed by helices α4 and α4' on one side (hereafter named the front entrance), and by helices α1, α1', and the two DNA recognition helices α3 and α3' on the other side (back entrance). The inside of the pore is largely hydrophobic and formed by residues from the N- and C-terminal helices of both subunits. The pore centre is constricted by a dyad-related pair of tryptophan residues (W96 in α4 and α4'), whose indole rings are oriented face-to-face at a distance of about 7 Å (as calculated from the centres of mass of the indole groups). Clusters of arginine and lysine residues surround the back entrance (K6, R10, K55, R59, R75 and K77 from each subunit) resulting in a net positive surface charge, which is consistent with that side of the dimer forming the binding site for DNA. By contrast, the surface around the front entrance is largely negatively charged due to the presence of twelve Glu and Asp residues, six from each C-terminal helix in the dimer (E83, E87, E94, D100, E104 and E107). The opposite electrostatic surface charges around the front and back entrances (Figure 1B) create a small but significant overall molecular dipole moment (1192 Debye) running through the pore coinciding with the dyad axis.

The dimer interface

A striking feature of the LmrR structure is its unusual dimeric arrangement leading to the formation of a large central pore. As pointed out above the dimer interface is formed by interactions of the C-terminal helix of each subunit with the DNA-binding domain of its dimer mate. The surface area of one subunit that becomes buried upon dimerization is ~1160 $Å^2$, which is within the expected range for a stable dimer considering the 13.5 kDa size of the LmrR subunit (Janin et al, 1988). However, the buried surface area in the LmrR dimer is substantially smaller than the buried surface areas in other MarR/PadR dimers (*e.g.* the buried surface area in the MarR dimer is 3700 $Å^2$), in which dimerization usually involves more inter-helical interactions and the

formation of a central compact core. Stabilization of the LmrR dimer mainly occurs via hydrophobic interactions. Residues in helix α4 that participate in forming the dimer interface are L91, A92, W96, R98, V99, I102, I103, N105 and L106. In the wHTH domain of the dimer-related subunit the residues important for dimerization are M8', A11', Q12', V15', I16' and V20' (from helix α1'), and V35', A38', N40', and M43' (from helix α2' and the loop connecting helices α2' and α3'). A few interactions are of polar or ionic nature. For example the side chain of N105 forms a hydrogen bond with the main chain oxygen of A38', and a salt bridge is formed between R98 and E42'. Another notable residue at the dimer interface is Q12' in helix α1'. The side chain of this residue forms a hydrogen bond with S95 in helix α4, and makes a π-cation interaction with W96 at the back face of the indole rings (assigning the one exposed towards the pore as their front face), thus stabilizing the conformation of this central residue in the pore.

Comparison with other winged helix proteins

A search of the PDB using the Dali server (Holm & Sander, 1996) showed that the LmrR subunit has significant structural homology to various DNA binding proteins containing helix-turn-helix or winged helix-turn-helix domains. Among these are the structurally and functionally characterized MarR/PadR family members MarR of *E. coli* (PDB accession code 1JGS, 87 equivalent Cα-atoms were superimposed with a root-mean-square-deviation (RMSD) of 2.3 Å), OhrR of *B. subtilis* (1Z9C, 86 equivalent Cα-atoms were superimposed with a RMSD of 2.8 Å), AphA of *Vibrio cholerae* (1YG2, 82 equivalent Cα-atoms were superimposed with a RMSD of 3.7 Å) and Pex of *Synechococcus sp.* (2E1N, 78 equivalent Cα-atoms were superimposed with a RMSD of 1.8 Å). The structural similarities of LmrR with the MarR/PadR proteins are mainly confined to its wHTH domain and the sequence identities are rather low (ranging from 11% for MarR to 28% for Pex). Interestingly, structural similarity was also detected with a number of hypothetical transcriptional regulators in the PDB, for which structures and function have not been published. One such protein, from *Clostridium thermocellum,* showed a particularly high structural homology with LmrR (PDB accession code 1XMA, 99 equivalent Cα-atoms were superimposed with a RMSD of 2.7 Å and a sequence identity of 35%). A structure-based sequence alignment of LmrR with 1XMA, AphA, MarR and OhrR is presented

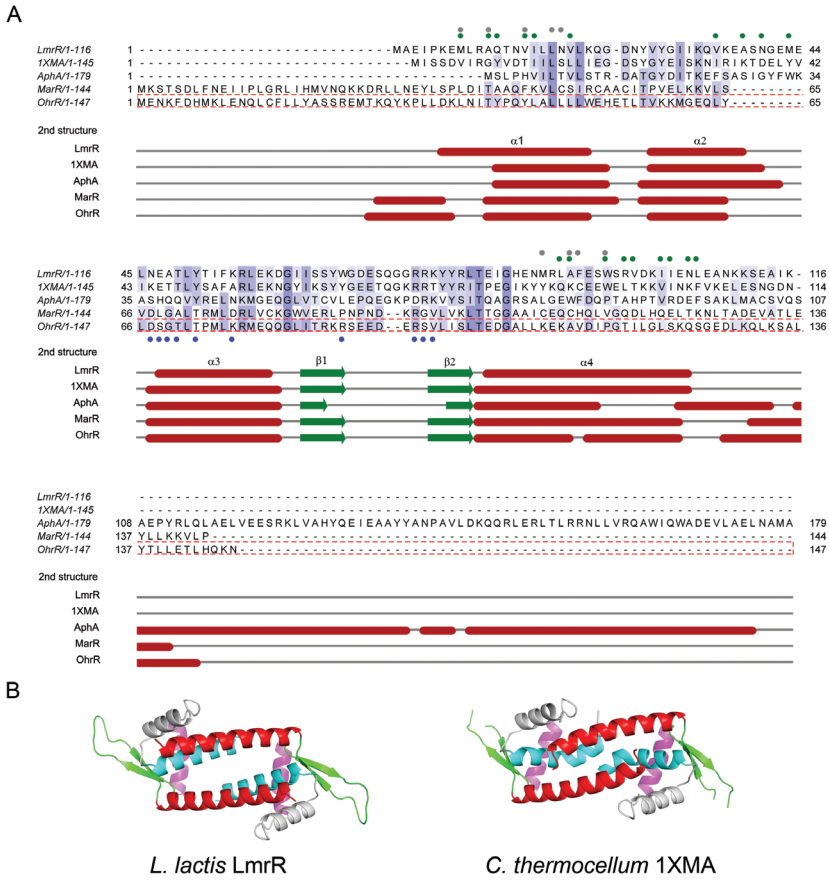

Figure 2. Comparison of LmrR with other wHTH proteins. (A) Structure-based multiple sequence alignment computed with the Tcoffee web server (Poirot et al, 2003) and visualized with JalView (Clamp et al, 2004). Structures were taken from the PDB using the following accession numbers: 1XMA, 1YG2 (AphA), 1JGS (MarR) and 1Z9C (OhrR). Residues of LmrR involved in dimerization are indicated with green dots, residues involved in drug binding with grey dots, and residues in OhrR involved in DNA binding are indicated with blue dots. (B) Ribbon representations of the LmrR and 1XMA dimers. Equivalent secondary structure elements are indicated in specific colors: α1 blue, α2 grey, α3 magenta, α4 red, and the β-wing green.

in Figure 2A. Most of the conserved residues are hydrophobic and appear to be important in stabilization of the overall fold of the DNA binding domain. Currently, the only MarR/PadR protein for which a DNA-bound structure is known is OhrR (Hong et al, 2005). The regions in LmrR that are equivalent to the DNA binding site in OhrR, *i.e.* helix α3 and the β-wing, show the highest

degree of conservation. Nevertheless, among the conserved residues only a few have a role in specific DNA binding in OhrR, suggesting that LmrR and OhrR recognize different DNA sequences. Interestingly, the putative DNA-binding sites of LmrR and 1XMA are highly conserved, which may indicate that these proteins bind similar DNA sequences. Unfortunately, the specific DNA binding sequence(s) of LmrR within the *lmrR* and *lmrCD* promoters have not yet been delineated (Agustiandari et al, 2008), and for the *C. thermocellum* homolog the target promoters are unknown. The comparison of LmrR with the four MarR/PadR transcription factors further reveals two invariant residues, G61 and G85, which are found in regions that connect the β-wing to helices α3 and α4 and appear to have crucial structural roles in stabilizing the conformation of the β-wing relative to the DNA recognition helix and the dimerization domain. Another invariant residue is T82, which is located at the proximal end of the wing between β2 and helix α4. Previously, it was shown that a T82I mutation in LmrR is associated with drug-resistant phenotypes of *L. lactis* (Lubelski et al, 2006), and that the LmrR-T82I mutant is deficient in both drug and DNA binding (Agustiandari et al, 2008). In the LmrR structure the side chain hydroxyl group of T82 is making three hydrogen bonds, with the backbone carbonyl of G61 at the N-terminal end of the β-wing and with the backbone amides of I84 and G85 at the N-terminus of helix α4, while the main chain amide of T82 forms a hydrogen bond with the backbone carbonyl of I62 (Supplementary Figure S2). These interactions are conserved in the other MarR/PadR structures, and appear to clamp the β-wing in place and stabilize the C-terminal helix. Replacement of T82 by an isoleucine residue will result in the loss of at least three of the four hydrogen bonds, as well as in the introduction of significant steric strain due to the presence of a bulkier and more hydrophobic side chain. The T82I mutation thus is predicted to induce a structural perturbation of LmrR, which probably corresponds to a detachment of the β-wing and a disordering of its C-terminal helix. Such a structural perturbation would explain the deleterious effect of the T82I mutation on overall LmrR function.

The structural organization of the LmrR dimer, with the two DNA-binding domains adjacent to one another and their DNA recognition helices facing the same side, is like in other MarR/PadR family members, pointing to a similar mode of DNA binding with the recognition helices fitting into two successive major grooves on one side of the DNA double helix. However, the dimeric architecture of LmrR is unique: none of the other structurally characterized

MarR/PadR family members shows a central pore at the dimer interface. It should be noted, though, that the dimerization modules in MarR/PadR proteins are highly diverse, consisting of different numbers of α-helices packed together in various ways. Of the four proteins that were compared with LmrR, only 1XMA has an identical overall topology containing a single C-terminal dimerization helix. Interestingly, the sequence homology of LmrR with 1XMA extends into the C-terminal helix and among the conserved residues are W96, as well as residues near the C-terminus that in LmrR participate in dimer formation. Nevertheless, unlike in LmrR, in the *C. thermocellum* transcription factor the N- and C-terminal helices of the two subunits form a compact core at the dimeric interface that is completely closed. This difference at the dimer interface is coupled to a difference in the relative orientation of the C-terminal helices with respect to the wHTH domains and to significant bending of the N- and C-terminal helices (Figure 2B), which allow the subunits in the dimer of the *C. thermocellum* transcription factor to approach each other more closely than the subunits in the LmrR dimer. To predict whether the "closed" conformation of the 1XMA dimer would be accessible to LmrR, a 1XMA-based homology model was prepared of LmrR using the SWISS model server (http://swissmodel.expasy.org/). The results indicate that there are no major steric clashes that would prevent LmrR to adopt a "closed" conformation. However, in comparison with the 1XMA dimer, the conserved tryptophan pair and their surrounding residues in LmrR seem unsuitable to provide the necessary hydrophobic packing and interactions to stabilize a closure of the central pore (Supplementary Figure S3). On the other hand, additional overall conformation changes and rearrangements of side chains could perhaps create a well-packed hydrophobic core. Clearly, homology modeling alone is not sufficient to assess the likelihood of LmrR also adopting a "closed" conformation. More sophisticated methods, like molecular dynamic simulations or NMR, could perhaps provide an answer, but these are beyond the scope of the present study.

Binding of Hoechst 33342 and Daunomycin

The structures of LmrR bound with H33342 and daunomycin reveal that the central pore in the LmrR dimer serves as a multi-drug binding site (Figure 3). In both complexes the pore accommodates a single drug molecule, consistent

with the 1:2 (drug:LmrR subunit) stoichiometry of drug-binding obtained previously from fluorescent titrations with H33342 (Agustiandari et al, 2008). The two drugs show a common mode of binding: their flat ring systems are wedged in between the W96 and W96' side chains forming aromatic stacking interactions with each of the two indole systems (Figures 3C-D). Further stabilization of the bound drugs is provided by various hydrophobic contacts in the pore. Remarkably, no hydrogen bonds are observed between the

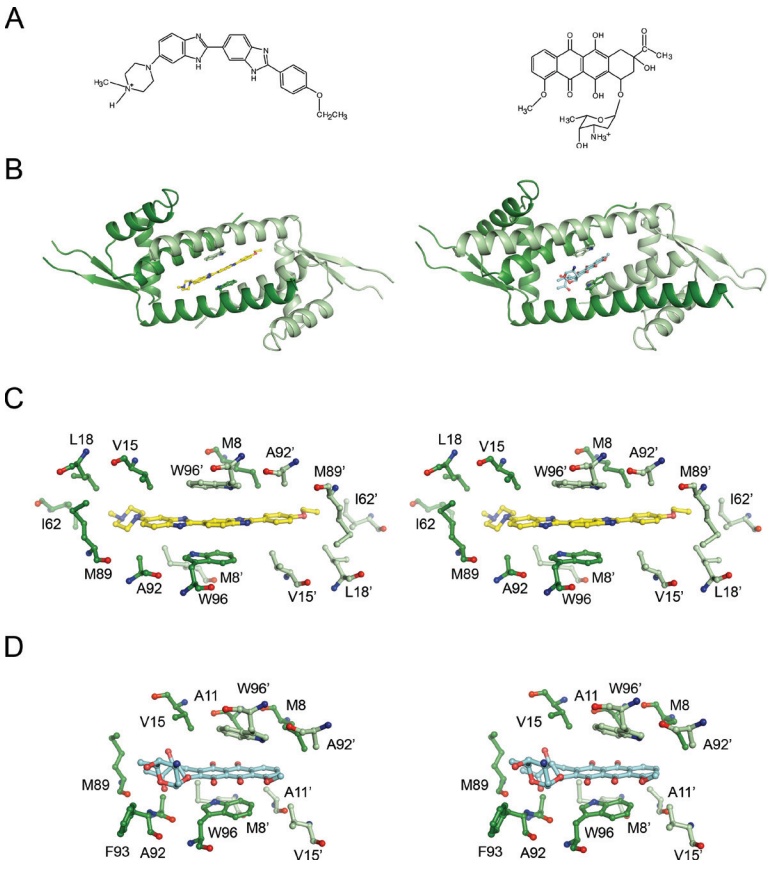

Figure 3. Crystallographic analysis of drug binding to LmrR. (A) Chemical structures of H33342 (left) and daunomycin (right). (B) Ribbon diagrams of the LmrR-H33342 (left) and LmrR-daunomycin (right) complex, showing the drug molecule (sticks) bound inside the central pore of the dimer in between the two tryptophan residues W96 and W96' (sticks). The two subunits of the LmrR dimer, as well as W96 and W96' are colored with different shades of green. (C) Close-up stereo view of the drug-binding pore in the LmrR-H33342 complex depicting the drug and residues that contact the drug (maximum contact distance defined as 4 Å). (D) Similar stereo view as in C of the drug-binding pore in the LmrR-daunomycin complex.

protein and the drugs. Rather, the orientation of the drugs is such that most of their polar atoms are in a solvent exposed position facing the front or back entrances of the pore. Several bound water molecules are present in the pore to act as hydrogen bond partners of the drugs.

The two drug-bound LmrR structures also show substantial differences. The elongated crescent-shaped Hoechst compound deeply penetrates the

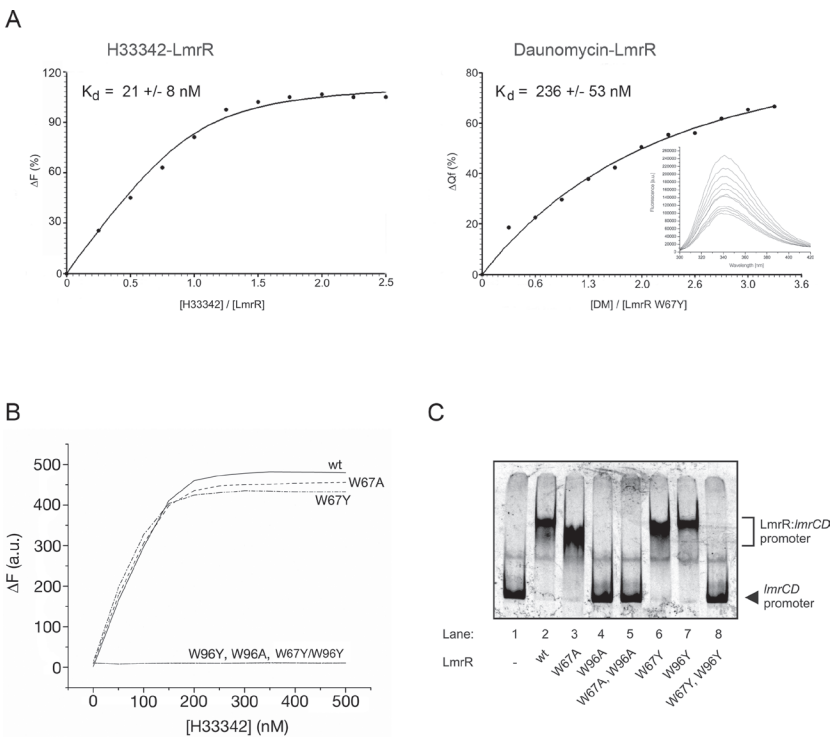

Figure 4. Spectroscopic and mutational analysis of drug and DNA binding by LmrR. (A) Fluorescence titration curves measuring H33342 and daunomycin binding to LmrR. Binding of H33342 was monitored by recording the increase in drug fluorescence when titrating a solution of wild-type LmrR with increasing concentrations of H33342, using experimental conditions as described (Agustiandari, 2008). Binding of daunomycin was monitored by performing tryptophan fluorescence quenching titration of the LmrR mutant W67Y with increasing concentrations of daunomycin. The apparent K_D values are calculated from the fitting of the data using non-linear regression analysis (r^2=0.993 and 0.996 for H33342 and daunomycin, respectively). Inset, tryptophan-fluorescence-emission spectra in the presence of increasing concentrations of daunomycin. (B) Fluorescence titration curves measuring H33342 binding to the W67 and W97 single mutants and the W67Y/W96Y double mutant of LmrR. The binding curves for the W96 single mutants, as well as the W67A/W96Y double mutant, are flat, thus revealing that these mutants lack drug-binding capability. (C) Electrophoretic mobility shift assay (EMSA) of the LmrR mutants analyzing their binding to the *lmrCD* promotor DNA.

pore and stretches out over its entire width with the ethoxy-phenolic and N-methyl-piperazine groups extending towards the sidewalls and facing the back entrance of the pore. The area of the drug-protein interaction surface is substantial (~215 Å^2) and shows good shape-complementarity. W96 and W96' clamp down one of the central benzimidazole ring systems, which is aligned in an off-centered parallel orientation with respect to each of the two indole rings. The inter-ring distances are ~3.5 Å, optimal for allowing the formation of strong van der Waals interactions (McGaughey et al, 1998)

Less extensive interactions are formed with daunomycin. The presence of the bulky amino sugar substituent prohibits a deep penetration of the drug into the flat pore (Figure 5C). Only the aglycon chromophore interacts with the protein, while the amino sugar is exposed to the solvent at the front entrance of the pore. As a consequence the area of the drug-protein interaction surface in the LmrR-daunomycin complex is much smaller (~82 Å^2) than in the LmrR-H33342 complex. No electron density could be observed for the amino sugar of daunomycin, indicating that this substituent is highly flexible. The stacking interactions of the aromatic rings with the W96/W96' pair form the main contribution to drug binding stabilization, but the stacking geometry is less optimal than in the LmrR-H33342 complex. The weaker interactions between LmrR and daunomycin, as compared to the interactions between LmrR and H33342, are further illustrated by the relatively high atomic B-factors of daunomycin (Table 1).

Drug Binding Affinities and Importance of W96 for Drug and DNA Binding

To quantify the difference in binding affinity of LmrR for H33342 and daunomycin the dissociation constants of the two drugs were approximated from binding curves obtained from two different, fluorescence-based drug-binding assays (Figure 4A). Binding of H33342 to untagged LmrR was monitored by recording the increase in drug fluorescence when H33342 moves from an aqueous to a hydrophobic environment, *i.e.* when it binds to the drug binding site of LmrR. The apparent dissociation constant of H33342 obtained from fitting the binding curve was ~20 nM, showing that this compound has a strong affinity for the drug binding site of LmrR. Unfortunately, the spectral properties of daunomycin did not allow the use of a similar binding assay. Instead, for

daunomycin, we obtained a binding curve by measuring the fluorescence quenching of W96 upon titration of the drug. LmrR contains two tryptophan residues, W67 and W96, the former of which is located in the β-wing of the DNA binding domain. To avoid unwanted disturbances of the fluorescence signal, W67 was mutated to either alanine or tyrosine. The W67Y and W67A mutations had no significant effect on the binding of H33342 by LmrR, nor on the binding of *lmrCD* promoter DNA (Figures 4B and 4C). The dissociation constant of daunomycin, determined from W96-fluorescence quenching using the LmrR-W67Y mutant, was ~0.25 μM, confirming that its binding affinity for LmrR is weaker than that of the Hoechst compound.

To confirm its importance for drug binding, W96 was also mutated to alanine or tyrosine. Both these LmrR mutants, W96A and W96Y, lost the ability to bind H33342 (Figure 4B). Interestingly, although the W96Y mutant was still able to bind to the *lmrCD* promotor, the W96A mutant was not, nor was a W67Y/W96Y double mutant (Figure 4C). The impaired DNA binding of the W67Y/W96Y mutant is not due to a loss of structural integrity of the protein, as was assessed by circular dichroism (Supplementary Figure S4A). These results thus point to some indirect role in DNA binding for residue 96, in addition to a direct role in drug binding.

Conformational flexibility

As pointed out earlier notable conformational differences were identified between the different LmrR structures (Figure 5). Pair-wise structural superpositions of the isolated subunits in the different dimers, using all Cα atoms, result in root-mean-square deviations ranging from 1.0 to 2.0 Å. These deviations primarily result from differences in the orientation of the C-terminal helix α4 relative to the wHTH domain. The reorientations of helix α4 may be described as lever-arm rotations with the residues that attach the N-terminus of the helix to the wHTH domain serving as a hinge (Figure 5A).

Even though the rotations are small (varying between 9° and 16°), the lever-like movement results in substantial translational shifts of residues near the C-terminus of helix α4 (up to ~8 Å as calculated from Cα-Cα distances). Except for small shifts of helices α1 and α2 no significant conformational changes are observed in the wHTH domains. Via the dimeric interface the conformational differences in the subunits are coupled to two pronounced

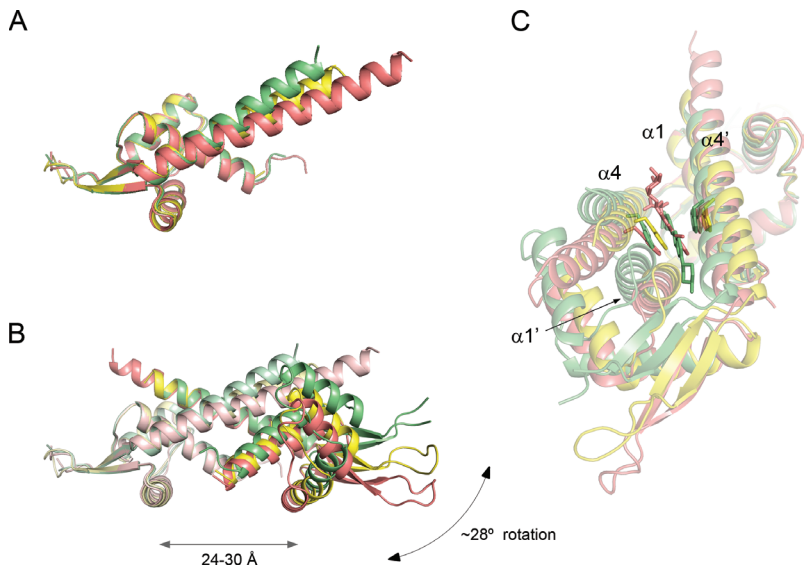

Figure 5. Conformational differences between the apo and two drug-bound structures of LmrR. (A) Superposition (in ribbon representation) of the apo-LmrR subunit structure (yellow) and the subunit structures of H33342-bound (green) and daunomycin-bound LmrR (salmon). The superposition was carried out using the Ca-atoms of the wHTH domain. (B) Superposition of the three LmrR dimers, showing the difference in relative position of the two wHTH domains. Only one of the two subunits (light colors) was used for the superposition (identical to the superposition in Figure 5A). Indicated are the range of distances between the two DNA recognition helices in the different dimers, and the largest rotational shift of the wHTH domains (based on comparing the H33342-bound and daunomycin-bound dimers). (C) The same superposition as in Figure 5B, but from a different view, showing the relative shifts of helix pair a1-a4' with respect to helix pair a1'-a4. Also shown in sticks are H33342 (green) and daunomycin (salmon), as well as the W96/W96' tryptophan pair. The sugar moiety of daunomycin is shown in a solvent exposed position at the front entrance of the pore, but it should be noted that its binding is highly disordered, as evident from the weak electron density associated with this substituent.

differences in the LmrR dimeric structures. Firstly, when comparing the H33342-bound and daunomycin-bound LmrR dimers with the apo-LmrR dimer, one of the wHTH domains is rotated relative to the wHTH domain of the other subunit. The amounts of rotation are about equal for both drug-bound complexes (~14°), but the rotations are in opposite directions (Figure 5B). In fact, the largest conformational change (~28° rotation) is observed when mutually comparing the H33342-bound and daunomycin-bound LmrR dimers. The rotations of the wHTH domains are coupled to changes in the spacing between the two DNA recognition helices. The shortest spacing is observed

in the daunomycin bound complex with a distance of ~26 Å (measured from the centroids of the helices), while in the H33342 bound complex the spacing is the largest with a distance of ~32 Å. Secondly, inside the pore differences are observed in the orientations of residues from one subunit relative to those of the other subunit. These differences are caused by shifts of the α1-α4' helix pair relatively to the α1'-α4 helix pair in the different LmrR structures, and directly affect the geometry of the central drug-binding site (Figure 5C).

Unfortunately, since each crystal structure of LmrR represents a different crystal form, it is not possible to distinguish whether the conformational changes are drug-induced, or whether they are caused by differences in crystal packing. However, the observed structural differences point to a remarkable plasticity of LmrR. As the conformational rearrangements affect both the drug binding and DNA binding sites in the LmrR dimer, similar conformational changes likely play an important role in the induction mechanism of LmrR.

DISCUSSION

The crystal structure of LmrR is the first structure of a transcription factor regulating the expression of a multi-drug ABC transporter. It is also the first time that the structural basis of multi-drug recognition has been studied for a PadR transcriptional regulator. Current knowledge of the mechanism of multi-drug recognition by transcriptional regulators of multi-drug transporters is largely based on crystallographic studies with the transcription factors BmrR from *Bacillus subtilis* (Zheleznova et al, 1999), QacR from *Staphyloccoccus aureus* (Schumacher & Brennan, 2003; Schumacher et al, 2001) and TtgR from *Pseudomonas putida* (Alguel et al, 2007). The structures of LmrR bound to H33342 and daunomycin confirm the importance of several of the general structural features that seem important in multi-drug recognition, *e.g.* the availability of a large drug-binding pocket that can accommodate a large spectrum of drug ligands, the importance of aromatic and hydrophobic residues for providing van der Waals interactions to stabilize the bound ligands, the importance of water molecules for occupying regions of the pocket not occupied by the ligand and for solvating hydrophilic groups of the ligand that do not interact with the protein, and a flexible pocket wall that can change conformation upon ligand binding. However, there are also some striking differences in the way multi-drug binding is accomplished by LmrR as compared to BmrR and

QacR. The foremost difference is that in LmrR the multi-drug binding pocket is formed by a symmetric pore located at the dimer centre with both subunits contributing equally to its architecture. In QacR and BmrR the drug-binding pockets are asymmetric and primarily formed within a single subunit. Also, in the drug-bound LmrR complexes the binding modes of the two different drugs are very similar and involve a common and strong aromatic stacking interaction with the W96/W96' tryptophan pair. In the other transcription factors the architecture of the drug-binding pockets allows different ligands to adopt different orientations within the pocket and to interact with different sets of amino acids. In BmrR and QacR, the binding affinity for cationic drugs is further augmented by electrostatic attraction between the positively charged ligand and buried or partially buried negatively charged glutamates or aspartate residues of the protein. No such interactions were observed in the H33342- and daunomycin-bound LmrR structures. However, in LmrR positive charges in the drugs may be stabilized by long-range electrostatic interactions with the cluster of glutamate and aspartate residues that surround the front entrance of the pore. Also the apparent molecular dipole moment that was found running through the pore of LmrR may assist in attracting and binding cationic drugs. The relevance of this latter feature of the LmrR structure for drug specificity is however unclear: it could be likewise important for directing LmrR towards the DNA substrate.

It is evident that the homology of LmrR with other members of the PadR protein family that have so far been functionally and structurally characterized, *i.e.* AphA and Pex, is relatively low. Our results thus confirm the classification of two PadR-subfamilies (Huillet et al, 2006): subfamily I with longer sequences (~180 amino acids) to which AphA and Pex belong, and a more distant subfamily II with shorter sequences (~110 amino acids) to which LmrR belongs. To the best of our knowledge LmrR is so far the only member of the PadR subfamily II that has been characterized both functionally and structurally. Interestingly, a BLAST search against translated nucleotide sequence databases (see Supplementary Figure S5) yields a large number of close homologs of LmrR that exist in various bacterial species, in particular those belonging to the Firmicutes (*Listeria, Bacillus, Staphylococcus, Enterococcus, Lactobacillus, Streptococcus, Clostridium*). Both the N- and C-terminal domain of LmrR are significantly conserved in these proteins, including W96, thus pointing to a high similarity in overall structure and dimeric organization.

Likely, some of these proteins also have a similar role as LmrR in regulating multi-drug resistance, but lack of functional data prohibits such assessment for the moment. Furthermore, the sequence conservation in the C-terminal helix of these LmrR-like proteins is no guarantee for the existence of a central multi-drug binding pore, as is evident from the structural comparison of LmrR with the *C. thermocellum* homolog. Future studies should therefore reveal whether the multi-drug binding characteristics of LmrR are applicable to a larger set of proteins.

In the absence of a DNA-bound structure, the induction mechanism of LmrR remains to be determined. Comparison of the different LmrR structures, and the effects of the W96 mutations, reveals a possible allosteric coupling between the drug and DNA binding sites. Most likely the binding of a drug to LmrR locks the dimer in a conformational state that is incompatible with DNA binding, due to a relative positioning of the DNA recognition helices that is unsuitable for their simultaneous insertion in the successive major grooves of the DNA. Such an induction mechanism would be similar to the induction mechanism of various wHTH-domain containing transcription factors, although the origin and nature of the structural changes involved in this mechanism are likely to be different for LmrR. To explore this further we tested the suitability of the different LmrR structures to bind B-form DNA using model building. None of the three LmrR structures has a conformation that allows a good fit with DNA (Supplementary Figure S6). In daunomycin-bound LmrR the DNA recognition helices are spaced too close together and would sterically clash with the DNA. In contrast, in the apo form and H33342-bound complex of LmrR they are positioned too far apart, and one of the two wHTH DNA binding domains is shifted away from the DNA. It should be noted, though, that such modeling does not take into account the possibility of DNA distortion, thus limiting its significance. Further studies are underway to better define the DNA binding characteristics of LmrR and unequivocally identify its operator DNA sequence. This knowledge will be crucial to allow crystallization of a DNA-bound LmrR complex.

MATERIALS AND METHODS

Chemicals

H33342 (Molecular Probes) and daunomycin (Calbiochem) were purchased and used without further purifications.

Protein production, crystallization and X-ray data collection

Tagged LmrR, comprising full length LmrR (from *L. lactis* strain MG1363) and a C-terminal streptactin-tag (117-SRWSHPQFEK-126), was obtained by nisin-induced overexpression in *L. lactis* using the expression vector pNSC8048-*lmrR*, and initially purified via strep-tag affinity and heparin column chromatography as described elsewhere (Agustiandari et al, 2008). Final purification to homogeneity was achieved by size exclusion chromatography on a Superdex 200 10/300 GL column (GE Healthcare) with a running buffer containing 20 mM Tris-HCl, pH 8.0, 280 mM NaCl and 1 mM EDTA. To anticipate possible negative effects of the streptactin-tag on protein crystal growth or protein conformation, also untagged full length LmrR was produced using a modified version of the overexpression plasmid in which the strep-tag encoding sequence was deleted. Untagged LmrR was first purified on a heparin column with conditions similar to those used for tagged LmrR, followed by purification on a Mono S HR 5/5 cation exchange column (GE Healthcare) with a linear gradient of 0.03-1 M NaCl (20 mM Hepes, pH 8, 1 mM EDTA, 0.5 mM DTT). Finally, the protein was loaded on a Superdex 200 10/300 GL column and eluted using the same running buffer as for tagged LmrR.

Purified LmrR, with or without streptactin-tag, was concentrated in the gel filtration running buffer and either used immediately for crystallization or frozen in liquid nitrogen and stored at -80 ^{o}C. Crystallization trials were set up with the aid of an Oryx6 crystallization robot (Douglas Instruments) using the PACT and JCSG crystallization screens (Newman et al, 2005). Lead conditions were optimized manually using the sitting drop vapor diffusion method with crystallization drops containing 1 µl of the protein solution (8 mg/ml) and 1 µl of the reservoir solution. Crystals of drug-bound LmrR complexes were obtained by co-crystallization using a LmrR solution (8 mg/ml) preincubated

for 30 minutes with 2 mM H33342 or daunomycin. Diffracting crystals of LmrR in drug-free conditions were obtained with both tagged and untagged protein, while in the drug-bound form only streptactin-tagged LmrR yielded well diffracting crystals. At drug-free conditions tagged LmrR crystals were obtained with 20% PEG 3350 in 0.1 M Bis-Tris propane, pH 8.5 and 0.2 M $NaNO_3$, while drug-free untagged LmrR crystals grew in 30% PEG 1500, 0.1 M propionic acid/cacodylate/Bis-Tris propane (PCB) cocktail buffer, pH 8.5. Crystals of H33342-bound LmrR grew against a well solution containing 25% PEG 1500, 0.1 M succinic acid/phosphate/glycine (SPG) buffer, pH 9.0, while crystals of daunomycin-bound LmrR-strep were obtained with 25% PEG 1500, 0.1 M malonic acid/imidazol/boric (MIB) buffer, pH 7.0. All crystals grew overnight at room temperature.

X-ray diffraction data were collected at cryogenic temperatures by using the MX beam lines of the European Synchrotron Radiation Facility (ESRF) at Grenoble. Prior to data collection, crystals were flash cooled in a cryoprotectant solution of mother liquor with 20% glycerol. The data were processed with MOSFLM (Leslie, 2006) and merged using SCALA as implemented in CCP4 (Winn et al, 2011). Relevant data statistics are shown in Table 1. Data on the tagged LmrR crystal grown in the absence of drugs, and its derived structure, will not be presented here, as they were merely used to verify that the C-terminal streptactin-tag did not affect the overall LmrR structure.

Structure determination of apo-LmrR

In the absence of drugs untagged LmrR crystallized in space group $C222_1$ with two subunits from two different dimers in the asymmetric unit. The crystals diffracted up to 2.0 Å resolution. The untagged apo-LmrR structure was solved by molecular replacement using PHASER with the automated search process (McCoy, 2007). Various search models were prepared and tried using the structures of homologous proteins from the PDB, as identified by the FFAS server (Jaroszewski et al, 2005). Molecular replacement succeeded with a search ensemble containing the structures of three hypothetical transcription factors (PDB entries 2ESH, 1YYV and 1XMA) having sequence identities with LmrR ranging from 15% to 34%. Phase improvement and construction of an initial protein model was performed by using the automatic map improvement and model building routines in RESOLVE (Terwilliger, 2003). The final model

Table I. Selected Crystallographic Data and Statistics

		LmrR-drug complexes	
	Apo-LmrR	H33342	Daunomycin
Data analysis			
Space group	$C222_1$	$P4_32_12$	$P2_12_12_1$
Unit cell (Å)			
a	46.6	34.9	35.4
b	52.6	34.9	53.0
c	174.9	197.0	147.1
Resolution (Å)	35 - 2.0	65 - 2.2	70 - 2.2
R_{merge} [a]	0.04 (0.5)[b]	0.05 (0.206)	0.036 (0.273)
Mean I/ĴI	46.6 (2.0)	21.1 (4.8)	28.9 (3.8)
Completeness (%)	100 (99.3)	99.8 (99.0)	96.0 (99.9)
Unique reflections	15006	6948	17020
Redundancy	6.7 (7.8)	3.9 (2.9)	3.0 (3.5)
Refinement statistics			
Resolution (Å)	23 - 2.0	50 - 2.2	50 - 2.2
R_{work}/R_{free} [c]	0.21/0.26	0.20 / 0.25	0.23/0.27
No. of non-H atoms			
Protein	1679	850	1859
Ligand	-	34	38
Waters	83	27	116
Root-mean-square deviations in			
bond length (Å)	0.011	0.015	0.019
bond angles (°)	1.4	1.0	0.8
Average B-values (Å^2)			
Protein	15	36	27
Ligand	-	42	74
Ramachandran analysis			
Most favoured (%)	99.0	98.0	98.0
Additional allowed (%)	1.0	2.0	2.0

[a] $R_{merge} = \Sigma\Sigma|I_{hkl} - I_{hkl}(j)|/ \Sigma I_{hkl}$, where $I_{hkl}(j)$ is the observed intensity and I_{hkl} is the final average intensity value.

[b] Values in parentheses are for the highest-resolution shell.

[c] $R_{work} = \Sigma||F_{obs}| - |F_{calc}||/\Sigma|F_{obs}|$ and $R_{free} = \Sigma||F_{obs}| - |F_{calc}||/\Sigma|F_{obs}|$, where all reflections belong to a test set of 10% randomly selected data.

was obtained by carrying out various cycles of refinement using REFMAC (Murshudov et al, 1997) interspersed with cycles of rebuilding and placement of water molecules using COOT (Emsley & Cowtan, 2004). TLS refinement was used in the last refinement cycles to model anisotropic displacements (Winn et al, 2001; Winn et al, 2003). The final model of the apo-LmrR structure contains two polypeptide chains: one discontinuous chain covering residues 5-70 and 75-109, and one continuous chain covering residues 5-109. In both polypeptides residues 1-4 and 110-116 are missing due to weak or absent electron density. Each polypeptide is one subunit of different biological dimers that are formed by crystallographic 2-fold axes.

Structure determination of drug-bound LmrR

H33342-bound LmrR crystals belong to space group $P4_32_12$ with one monomeric molecule per asymmetric unit, while daunomycin-bound LmrR crystallized in a different space group ($P2_12_12_1$) with a dimer in the asymmetric unit. The structures of the drug-bound complexes were solved by molecular replacement using the apo-LmrR monomer as a search model. Clear density in $2F_o$-F_c and F_o-F_c Fourier maps, calculated at the initial stages of refinement, indicated the location and binding mode of the drugs. The model building and refinement were done with COOT and REFMAC5. In both cases TLS refinement was used in the last refinement cycles. For the LmrR-H33342 complex the final protein model contains residues 3-108. No electron density is observed for residues 1-2, 109-126 (including the strep-tag) and 71-73 (β-wing) loop. The final protein model of LmrR-daunomycin contains residues 2-116 for chain A, and residues 5-115 for chain B.

Structure analysis

Relevant crystallographic statistics of the refined models are shown in Table 1. Stereochemistry of the models was validated with the programs Procheck (Laskowski et al, 1993) and MolProbity (Davis et al, 2007). 3D structural superpositions and assessment of conformational differences was carried out with the programs Lsqman (Kleywegt, 1999) and Dyndom (Hayward & Berendsen, 1998). Electrostatic surface potentials were calculated using APBS

(Baker et al, 2001) and visualized using PyMOL (Delano Scientific). The molecular dipole moment of the LmrR dimer was calculated using the protein dipole moments server at http://bioportal.weizmann.ac.il/dipol. Additional analyses, like calculation of surface areas, were performed with various programs from the CCP4 program suite (Win et al, 2011).

Site directed mutagenesis

Mutations of W67 and W96 in the *lmrR* gene were performed via round PCR using the pNSC8048-lmrR plasmid as the template together with synthetic primers containing the designated mutations. PCR products were ligated at 4˚C overnight before being transformed to *L. lactis* NZ9000 competent cells via electroporation. Selected colonies were inoculated at 30˚C in M17 media (Difco) supplemented with 0.5% glucose (w/v) and 5 µg/ml chloroamphenicol. Plasmid isolation was performed using GenElute Plasmid miniprep kit (Sigma-Aldrich) and the correct mutations were verified by nucleotide sequencing.

Drug binding assays

Binding of H33342 to the DNA-free purified LmrR variants was monitored by the increase of H33342 fluorescence upon binding as described previously (Agustiandari et al, 2008). Binding of daunomycin to LmrR mutant W67Y was monitored by tryptophan fluorescence quenching titration experiments using an Aminco Bowman Series 2 spectrofluorometer (excitation wavelength of 295 nm, emission spectra obtained from 300 to 450 nm). A detailed description of the drug binding assays, additional control experiments, and the procedure that was followed to derive the apparent dissociation constants (K_d) of the two drugs is included in the Supplementary Information.

DNA binding assay

The ability of the LmrR variants to bind to a 287 bp fragment corresponding to the promoter region of *lmrCD* was studied by means of an electrophoretic mobility shift assay (EMSA) as described (Agustiandari et al, 2008).

Circular dichroism

Circular dichroism spectra were obtained at 25 °C by using an Aviv 62ADS spectropolarimeter (Aviv Associates, Lakewood, NJ). The protein samples contained 0.24 mg/ml protein in 20 mM Tris-HCl, pH 8.0, and 50 mM NaCl.

Accession Numbers

The atomic coordinates and structure factors for apo-LmrR (entry 3F8B), H33342-bound LmrR (entry 3F8C), and daunomycin-bound LmrR (entry 3F8F) have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University (http://www.rcsb.org).

Acknowledgements

We acknowledge the European Synchrotron Radiation Facility for provision of beam time and we thank the MX beamline scientists for assistance in beam line usage.

REFERENCES

Agustiandari H, Lubelski J, van den Berg van Saparoea HB, Kuipers OP, Driessen AJ (2008) LmrR is a transcriptional repressor of expression of the multidrug ABC transporter LmrCD in Lactococcus lactis. *J Bacteriol* **190**(2)**:** 759-763

Alekshun MN, Levy SB, Mealy TR, Seaton BA, Head JF (2001) The crystal structure of MarR, a regulator of multiple antibiotic resistance, at 2.3 A resolution. *Nat Struct Biol* **8**(8)**:** 710-714

Aravind L, Anantharaman V, Balaji S, Babu MM, Iyer LM (2005) The many faces of the helix-turn-helix domain: transcription regulation and beyond. *FEMS Microbiol Rev* **29**(2)**:** 231-262

Arita K, Hashimoto H, Igari K, Akaboshi M, Kutsuna S, Sato M, Shimizu T (2007) Structural and biochemical characterization of a cyanobacterium circadian clock-modifier protein. *J Biol Chem* **282**(2)**:** 1128-1135

Clamp M, Cuff J, Searle SM, Barton GJ (2004) The Jalview Java alignment editor. *Bioinformatics* **20**(3)**:** 426-427

Davis IW, Leaver-Fay A, Chen VB, Block JN, Kapral GJ, Wang X, Murray LW, Arendall WB, 3rd, Snoeyink J, Richardson JS, Richardson DC (2007) MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. *Nucleic Acids Res* **35**(Web Server issue)**:** W375-383

De Silva RS, Kovacikova G, Lin W, Taylor RK, Skorupski K, Kull FJ (2005) Crystal structure of the virulence gene activator AphA from Vibrio cholerae reveals it is a novel member of the winged helix transcription factor superfamily. *J Biol Chem* **280**(14)**:** 13779-13783

Emsley P, Cowtan K (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* **60**(Pt 12 Pt 1)**:** 2126-2132

Gury J, Barthelmebs L, Tran NP, Divies C, Cavin JF (2004) Cloning, deletion, and characterization of PadR, the transcriptional repressor of the phenolic acid decarboxylase-encoding padA gene of Lactobacillus plantarum. *Appl Environ Microbiol* **70**(4)**:** 2146-2153

Hayward S, Berendsen HJ (1998) Systematic analysis of domain motions in proteins from conformational change: new results on citrate synthase and T4 lysozyme. *Proteins* **30**(2)**:** 144-154

Heldwein EE, Brennan RG (2001) Crystal structure of the transcription activator BmrR bound to DNA and a drug. *Nature* **409**(6818)**:** 378-382

Higgins CF (2007) Multiple molecular mechanisms for multidrug resistance transporters. *Nature* **446**(7137)**:** 749-757

Holm L, Sander C (1996) Mapping the protein universe. *Science* **273**(5275)**:** 595-603

Hong M, Fuangthong M, Helmann JD, Brennan RG (2005) Structure of an OhrR-ohrA operator complex reveals the DNA binding mechanism of the MarR family. *Mol Cell* **20**(1)**:** 131-141

Huillet E, Velge P, Vallaeys T, Pardon P (2006) LadR, a new PadR-related transcriptional regulator from Listeria monocytogenes, negatively regulates the expression of the multidrug efflux pump MdrL. *FEMS Microbiol Lett* **254**(1)**:** 87-94

Janin J, Miller S, Chothia C (1988) Surface, subunit interfaces and interior of oligomeric proteins. *J Mol Biol* **204**(1)**:** 155-164

Jaroszewski L, Rychlewski L, Li Z, Li W, Godzik A (2005) FFAS03: a server for profile--profile sequence alignments. *Nucleic Acids Res* **33**(Web Server issue)**:** W284-288

Kleywegt GJ (1999) Experimental assessment of differences between related protein crystal structures. *Acta Crystallogr D Biol Crystallogr* **55**(Pt 11)**:** 1878-1884

Laskowski RA, MacArthur MW, Moss DS, Thornton JM (1993) PROCHECK: a program to check the stereochemical quality of protein structures. *J Appl Cryst* **26:** 283-291

Leslie AG (2006) The integration of macromolecular diffraction data. *Acta Crystallogr D Biol Crystallogr* **62**(Pt 1)**:** 48-57

Lubelski J, de Jong A, van Merkerk R, Agustiandari H, Kuipers OP, Kok J, Driessen AJ (2006) LmrCD is a major multidrug resistance transporter in Lactococcus lactis. *Mol Microbiol* **61**(3)**:** 771-781

McCoy AJ (2007) Solving structures of protein complexes by molecular replacement with Phaser. *Acta Crystallogr D Biol Crystallogr* **63**(Pt 1)**:** 32-41

McGaughey GB, Gagne M, Rappe AK (1998) pi-Stacking interactions. Alive and well in proteins. *J Biol Chem* **273**(25)**:** 15458-15463

Murshudov GN, Vagin AA, Dodson EJ (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr D Biol Crystallogr* **53**(Pt 3)**:** 240-255

Newman J, Egan D, Walter TS, Meged R, Berry I, Ben Jelloul M, Sussman JL, Stuart DI, Perrakis A (2005) Towards rationalization of crystallization screening for small- to medium-sized academic laboratories: the PACT/JCSG+ strategy. *Acta Crystallogr D Biol Crystallogr* **61**(Pt 10)**:** 1426-1431

Poirot O, O'Toole E, Notredame C (2003) Tcoffee@igs: A web server for computing, evaluating and combining multiple sequence alignments. *Nucleic Acids Res* **31**(13)**:** 3503-3506

Saier MH, Jr., Paulsen IT, Sliwinski MK, Pao SS, Skurray RA, Nikaido H (1998) Evolutionary origins of multidrug and drug-specific efflux pumps in bacteria. *FASEB J* **12**(3)**:** 265-274

Schumacher MA, Brennan RG (2003) Deciphering the molecular basis of multidrug recognition: crystal structures of the Staphylococcus aureus multidrug binding transcription regulator QacR. *Res Microbiol* **154**(2)**:** 69-77

Schumacher MA, Miller MC, Grkovic S, Brown MH, Skurray RA, Brennan RG (2001) Structural mechanisms of QacR induction and multidrug recognition. *Science* **294**(5549)**:** 2158-2163

Terwilliger TC (2003) SOLVE and RESOLVE: automated structure solution and density modification. *Methods Enzymol* **374:** 22-37

Winn MD, Ballard CC, Cowtan KD, Dodson EJ, Emsley P, Evans PR, Keegan RM, Krissinel EB, Leslie AG, McCoy A, McNicholas SJ, Murshudov GN, Pannu NS, Potterton EA, Powell HR, Read RJ, Vagin A, Wilson KS (2011) Overview of the CCP4 suite and current developments. *Acta Crystallogr D Biol Crystallogr* **67**(Pt 4)**:** 235-242

Winn MD, Isupov MN, Murshudov GN (2001) Use of TLS parameters to model anisotropic displacements in macromolecular refinement. *Acta Crystallogr D Biol Crystallogr* **57**(Pt 1)**:** 122-133

Winn MD, Murshudov GN, Papiz MZ (2003) Macromolecular TLS refinement in REFMAC at moderate resolutions. *Methods Enzymol* **374:** 300-321

Zheleznova EE, Markham PN, Neyfakh AA, Brennan RG (1999) Structural basis of multidrug recognition by BmrR, a transcription activator of a multidrug transporter. *Cell* **96**(3)**:** 353-362

Zsila F, Bikadi Z, Simonyi M (2004) Induced circular dichroism spectra reveal binding of the antiinflammatory curcumin to human alpha1-acid glycoprotein. *Bioorg Med Chem* **12**(12)**:** 3239-3245

Supplementary materials and methods

Drug binding assays — Protein concentrations were calculated from the absorbance at 280 nm (A_{280}). The H33342-binding assay was performed as follows: to a solution of purified LmrR (0.2 μM as dimer) in 50 mM Tris-HCl, pH 7.0, increasing amounts of H33342 were added and the fluorescence was followed at excitation and emission wavelengths of 355 and 457 nm, respectively, with slit widths of 5 nm using a Perkin-Elmer LS 50B spectrofluorometer at room temperature. Fluorescence data were corrected as described previously (Agustiandari et al, 2008). To monitor the binding of daunomycin to LmrR the following procedure was employed. To a solution of LmrR (0.15 μM of dimeric species, 120 μl) in 50 mM Tris-HCl, pH 8.0, small volumes (1.2 μl) of a daunomycin stock solution were added in sequential steps. After each titration step the solution was mixed thoroughly and incubated for 2 minutes before measurements were made. Tryptophan fluorescence was recorded at room temperature with an Aminco Bowman Series 2 spectrofluorometer using a 0.2 ml cell and an excitation wavelength of 295 nm. Emission spectra were obtained from 300 to 450 nm. The excitation and emission slit widths were 4 nm. Each spectrum was scanned twice to obtain the final fluorescence emission spectra. Spectra were corrected for background emission and dilution effects.

Apparent dissociation constants (K_d) of the two drugs were obtained by data fitting using non-linear regression analysis (program LabFit), employing the following equation (Zsila et al, 2004):

$$RF = \frac{k}{2}\left(C_P + C_L + K_d - \sqrt{\left(C_P + C_L + K_d\right)^2 - 4C_P C_L} \right)$$

where RF is either the relative increase in fluorescence intensity (DF) of H33342 upon binding to LmrR (expressed as $(I-I_0)/(I_{max}-I_0)$, with I_0 the intensity of fluorescence in the absence of ligand, I the intensity of fluorescence upon addition of ligand and I_{max} the maximum intensity of fluorescence at saturation) or the tryptophan fluorescence quenching (DQ_F) by daunomycin at 340 nm (expressed as $(I_0-I)/I_0$, with I_0 the intensity of fluorescence in the absence of quencher and I the intensity of fluorescence upon addition of quencher). C_p and C_L are the total concentrations of protein and ligand (drug), respectively, and *k* is a constant.

Control experiments — As a control the drugs were titrated to a solution of the neutral tryptophan analog N-acetyl-tryptophanamide (0.4 mM) in 20 mM Tris-HCl, pH 8.0, 270 mM NaCl, and the fluorescence data were recorded using an identical setup as with the LmrR protein solution. No significant changes in fluorescence were recorded (Supplementary Figures S4B-C), thus showing that the changes in fluorescence recorded for drug titration in the presence of LmrR are protein-driven, and not merely the result of an intrinsic affinity of the drugs for tryptophan.

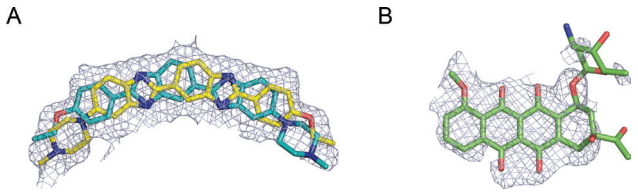

Figure S1 The binding of H33342 and daunomycin to LmrR. **(A)** σ_A-weighted *Fo-Fc* omit electron-density for bound H33342 (in sticks, showing the two binding modes that are related by the crystallographic dyad). Carbon atoms are shown in yellow or cyan (to distinguish the two binding modes), oxygen in red and nitrogen in blue. (**B**) σ_A-weighted *Fo-Fc* omit electron density for bound daunomycin. To remove model bias, the protein models (after removing the drugs) were subjected to 30 rounds of refinement by REFMAC (Murshudov et al, 1997), prior to map calculation. The contour level of the maps is 3σ (H33342) and 2.5σ (daunomycin) and the resolution is 2.2 Å.

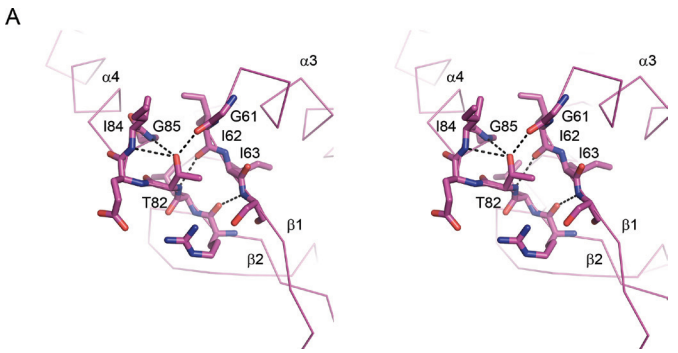

Figure S2 Stereo view showing the environment and interactions of T82 in the apo LmrR structure. Hydrogen bonds are indicated with dashed lines.

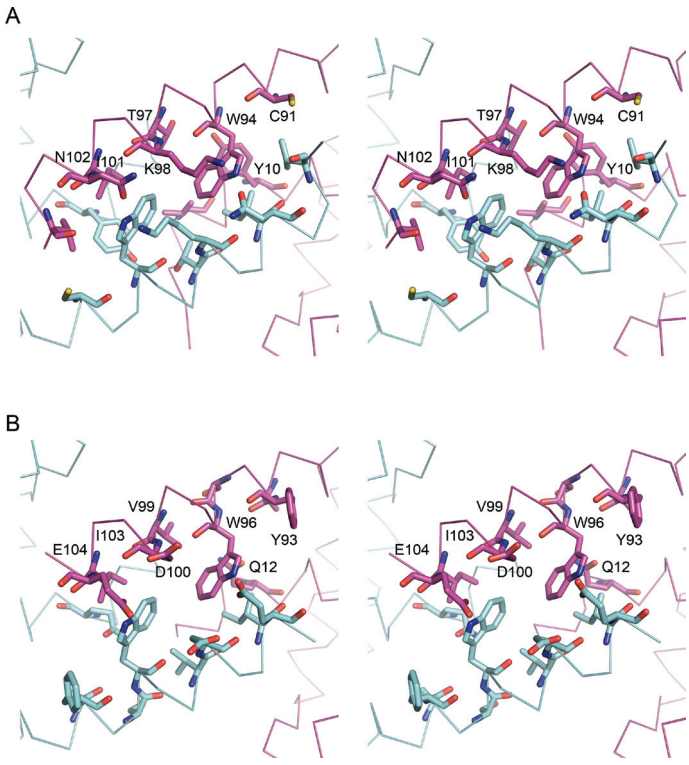

Figure S3 Assessment of LmrR adopting a 1XMA-like "closed" conformation. (**A**) Stereo view of the centre of the 1XMA dimer, along the 2-fold rotation axis, showing the environment of the conserved tryptophan pair (residues W94 and W94') and their close packing with neighbouring residues. The two subunits in the dimer are differently colored in magenta and cyan (oxygen atoms in red, nitrogen in blue, sulphur in yellow. Only residues of one subunit are labeled. (**B**) Similar stereo view as in A of the centre of a "closed" LmrR dimer, obtained via homology-modeling using the 1XMA dimer as a template. Labels refer to the residues that are equivalent to those in 1XMA.

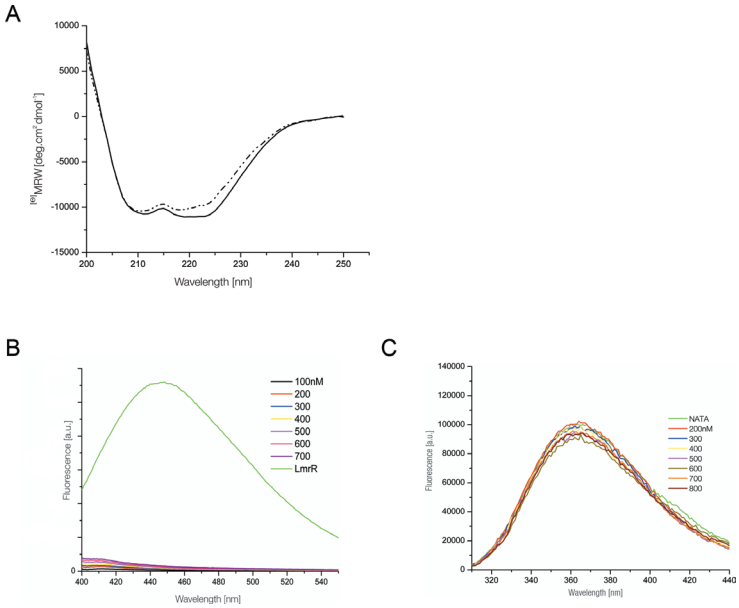

Figure S4 Spectroscopic control experiments. (**A**) Circular dichroism spectra of wild-type LmrR (solid line) and the LmrR-W67Y/W96Y double mutant (dashed line). (**B**) Fluorescence titration experiment of H33342 to a 0.4 mM solution of N-acetyl-tryptophanamide (NATA). The flat lines show the fluorescence spectra of the NATA solution after adding increasing amounts of H33342 (final concentrations ranging from 100 to 700 nM). The upper green line shows the fluorescence spectrum after adding LmrR (to a final concentration of 0.2 mM as dimer) to the H33342/NATA solution at the end of the titration series. (**C**) Tryptophan fluorescence spectra recorded after titrating increasing amounts of daunomycin (final concentrations 200-800 nM) to a 0.4 mM NATA solution.

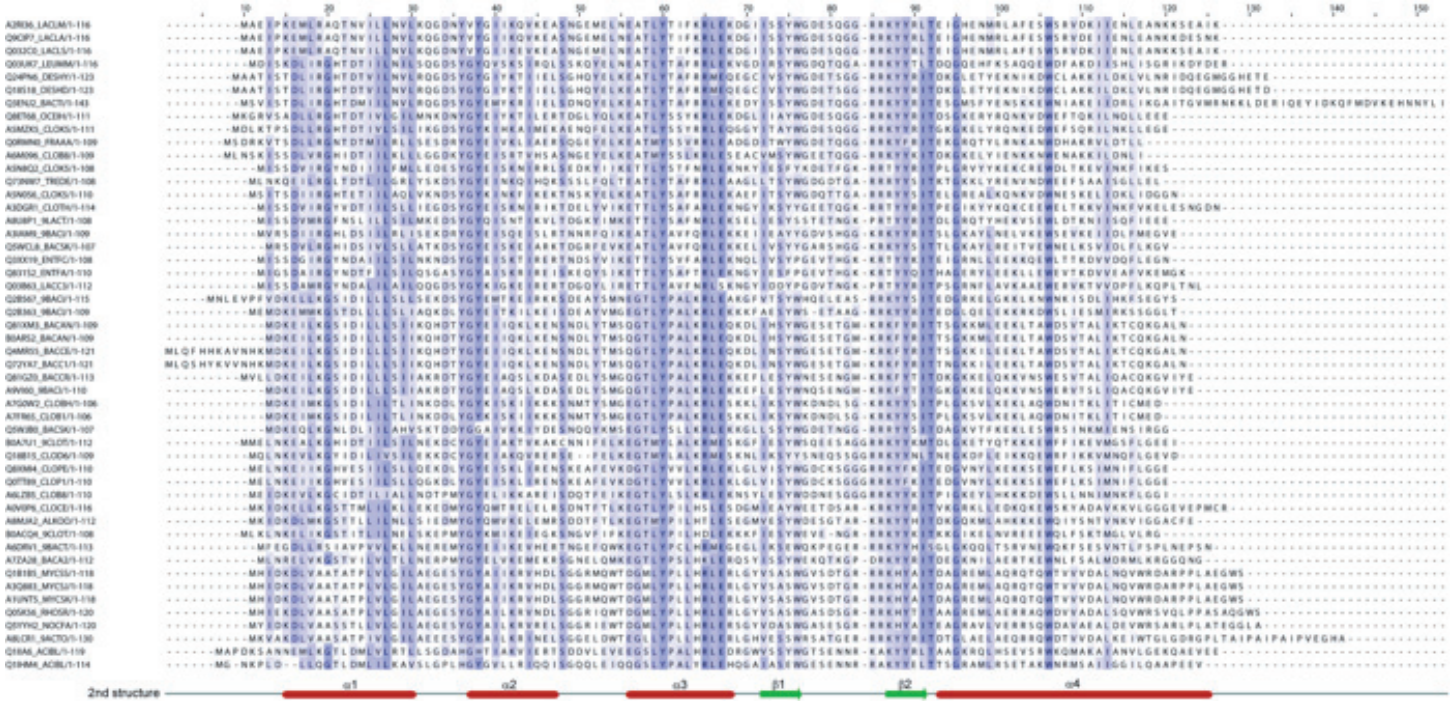

Figure S5 Multiple sequence alignment of LmrR from *L. lactis* strain MG1363 (trEMBL accession number A2RI36, top sequence) with a selected group of close homologs. Sequences were identified and aligned using BLAST, and the alignment figure was prepared with JalView (http://www.jalview.org). Sequences are abbreviated with their trEMBL entry names, including a reference to the source organism. The sequence identities with LmrR vary between 25-99% (the highest identity is with LmrR from *L. lactis* strain SK11) with 30% of the sequences having a sequence identity above 35%. The secondary structure assignment is based on the LmrR crystal structure. In blue boxes are the conserved residues (based on identities), with the more highly conserved residues indicated by darker shades.

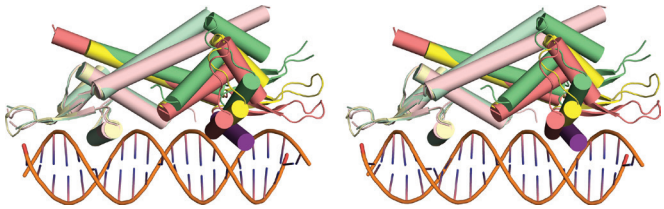

Figure S6 Stereo view of the three superimposed LmrR dimers (apo, H33342-bound and daunomycin-bound) docked to a 21 base-pair fragment of double-stranded B-form DNA. Docking was performed manually using the programs Coot (Emsley & Cowtan, 2004) and PyMOL (Delano Scientific). One of the two DNA recognition helices in each dimer (in the figure defined as the left one) could be optimally positioned in one of the two neighbouring major grooves of the DNA, but this resulted in a significant misalignment of the second DNA recognition helix with respect to the other major groove. The optimal location for the second DNA recognition helix that would allow the formation of binding interactions is indicated in dark magenta. Colors for the different LmrR structures are as follows: yellow, apo; green, H33342-bound; salmon, daunomycin-bound. The light and dark colors distinguish the two subunits in each dimer.

References

Zsila F, Bikadi Z, Simonyi M (2004) Induced circular dichroism spectra reveal binding of the antiinflammatory curcumin to human alpha1-acid glycoprotein. *Bioorg Med Chem* **12**: 3239-3245

Emsley P, Cowtan K (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* **60**: 2126-2132

Chapter

Multidrug resistance regulator LmrR uses a single site to bind structurally diverse compounds

3

Pramod Kumar Madoori, Abhishek Tomar, Arnold J. M. Driessen and Andy-Mark W. H. Thunnissen

Manuscript submitted

ABSTRACT

LmrR is a PadR-like transcription factor in *Lactococcus lactis* that regulates the drug-induced expression of the ABC-type multidrug efflux transporter LmrCD. Previously, we reported crystal structures of LmrR in an unliganded state, and in complex with the cytotoxic compounds Hoechst 33342 and daunomycin. Here, we reveal two additional crystal structures of LmrR bound to the structurally diverse compounds ethidium and riboflavin. The new structures confirm that LmrR uses a single ligand-binding site within a large, symmetrical and hydrophobic pore in the dimer interface to accomplish multidrug binding. Multidrug recognition is dominated by the aromatic stacking interactions of Trp96/Trp96' with the flat heterocyclic cores of the ligands. Additional drug binding stabilization is provided mainly by hydrophobic and van der Waals interactions with aliphatic amino acid residues. The hydrophilic ribityl moiety of riboflavin is less well accommodated within the hydrophobic binding site than the phenyl and ethyl moieties of ethidium, explaining the 20-fold lower binding affinity of the flavin compound. The measured binding affinity of riboflavin, a neutral ligand, was similar as the binding affinity measured previously for daunomycin, a cationic ligand, suggesting that electrostatic interactions with aspartate residues near the multidrug-binding site do not significantly contribute to ligand binding affinity. The use of a single site for binding the different ligands, rather than partially overlapping subsites, and the dominating role of the Trp96/Trp96' pincer in drug recognition and binding, distinguishes LmrR from other well-characterized multidrug-linked transcription regulators.

Introduction

When *Lactococcus lactis* is challenged with cytotoxic compounds like daunomycin, Hoechst 33342, ethidium and rhodamine 6G it readily displays a multidrug resistance (MDR) phenotype.[1] The major factor responsible for eliminating these toxic compounds from the cell is the ABC efflux transporter LmrCD.[2] Production of this transporter is controlled by the multidrug-sensitive transcriptional repressor LmrR.[2, 3] In the absence of drugs, LmrR prevents expression of the transporter by binding to the *lmrCD* promoter. When toxic substrates enter the cell, their binding to LmrR triggers an allosteric response, resulting in its release from DNA followed by derepression of *lmrCD* transcription and subsequent increase in multidrug efflux activity. Through an autoregulatory mechanism LmrR also represses transcription of its own gene, ensuring a fine-tuned demand-dependent expression of the LmrCD transporter.[4]

LmrR is a 116-residue protein that belongs to the PadR family of bacterial transcription regulators. Members of this large, but poorly characterized family have been implicated in the regulation of various adaptive responses in bacteria, leading to, for example, detoxification, MDR and virulence.[5,6] Apart from LmrR, LadR from *Listeria monocytogenes* is the only characterized member of the MDR-related PadR regulators.[7] Crystallographic analysis of LmrR in its drug-free state, and in complex with the substrates Hoechst 33342 (H33342) and daunomycin (DAU), has revealed the overall structural features of this protein and the location and architecture of its multidrug binding site.[8] LmrR forms a homodimer with each subunit containing a typical β-winged helix-turn-helix domain for DNA-binding, and a single C-terminal helix, which is responsible for dimerization. The overall structure of LmrR is reminiscent of that of members of the MarR/SlyA family of transcription factors,[9,10] which regulate similar biological processes in bacteria, including antibiotic resistance and virulence, but which generally have a much larger C-terminal dimerization domain. A highly unusual feature of LmrR is the presence of a symmetrical, flat-shaped hydrophobic pore at the dimer centre that contains the multidrug-binding site. H33342 and DAU bind in a similar manner with their aromatic rings sandwiched between the indole groups of two dimer-related tryptophan residues (W96/W96', the prime indicating the second subunit in the dimer).[8]

Multidrug-sensitive transcription regulators, like LmrR, are attractive targets for elucidating the mechanisms of multidrug recognition. They display similar substrate recognition profiles as the transporters they regulate, but, being soluble proteins, are more amenable to biochemical and structural studies. Our current understanding of multidrug binding and specificity by MDR-linked regulators is largely based on the crystallographic analysis of proteins belonging to the TetR and MerR family of transcription regulators, in particular QacR, TtgT, and BmrR.[11-17] Although they recognize a similar range of ligands, with a preference for cationic lipophilic compounds containing planar aromatic groups, these proteins use different binding site designs and mechanisms to accomplish multidrug recognition.[17,18] Unfortunately, the scope for getting insights into the mechanism of multidrug-recognition by LmrR, and how this mechanism relates to that used by QacR, TgtR and BmrR, is somewhat limited by the availability of only two structures of drug-bound LmrR complexes. Here we report two additional crystal structures of LmrR in complex with ethidium (ET) and riboflavin (RBF). The combined structural data of all four LmrR-drug complexes, together with solution-binding data, allows a better definition of the multidrug binding principles employed by LmrR, and demonstrates that its mechanism of multidrug recognition uniquely combines different features of other well-established multidrug binding models.

Materials and Methods

Crystallization and data collection

LmrR was produced as an untagged protein by nisin-induced overexpression in *Lactococcus lactis* and purified as described earlier.[8] Complexes of LmrR with riboflavin and ethidium were prepared by mixing protein and drug in a 1:2 molar ratio in a solution containing 20 mM Tris-HCl, pH 8.0, 280 mM NaCl and 1 mM EDTA. Initial crystallization conditions were obtained by sparse-matrix screening, using the PACT and JCSG+ commercial kits (Molecular Dimensions) and a Douglas Instruments Oryx-6 crystallization robot. Manual optimization using a sitting-drop vapor diffusion setup resulted in a final and equivalent crystallization solution for both protein-ligand complexes, containing 0.1 M Tris-HCl, pH 8.5, 17% PEG 2000 monomethyl ether (MME) and 0.2 M tri-

methylamine N-oxide. Crystals grew overnight from drops containing 1 µl reservoir and 1 µl of the protein-drug mixture at 295 K. To allow flash-cooling of the crystals, a cryo-protecting solution was prepared from the reservoir solution by increasing the concentration of PEG 2000 MME to 40%. Addition of 0.1 M NaCl and 0.2 mM ligand to the cryoprotecting solution was necessary to sufficiently stabilize the crystals. X-ray intensity data were collected from single flash-cooled crystals on beamline ID23-2 at the European Synchrotron Radiation Facility (ESRF) at 100 K. Data were processed with MOSFLM[19] or XDS[20] and scaled, merged, and reduced with programs from the CCP4 suit[32] (http://www.ccp4.ac.uk/).

Structure determination and refinement

Structures of LmrR complexes with riboflavin and ethidium were solved by molecular replacement with the program PHASER from the CCP4 program suite, using a single subunit of the unliganded LmrR dimer as a search model (PDB entry 3F8B). The models were improved in several cycles, by restrained refinement using the programs Phenix.refine[21] and Refmac5[22], alternated by manual model building using COOT.[23] Subsequently, the ligands were modeled into excess density observed in the central LmrR pore. Evaluation of the interaction geometries and real-space correlation factors was used to guide the docking of the ligands. Fixed translation libration screw (TLS) parameters were determined using the TLS motion detection server[24] and then used in the subsequent rounds of structure refinement. In the last stages of the refinements, water molecules were placed and retained in the model by strict criteria of difference density, B-factor cutoffs, and hydrogen-bonding capacity. The quality of the final models was checked using MolProbity.[25] Selected data collection and refinement statistics are presented in Table I. The coordinates and structure factors have been deposited in the RCSB with PDB accession codes 3UJY and 3UJ5 for LmrR•RBF and LmrR•ET, respectively.

Structure analysis

The surface and volume of the central pore were generated using CASTp.[26] Solvent-accessible buried surfaces and drug/binding site shape

complementarities were calculated using the programs Areaimol and Sc from the CCP4 program suite. Ligand interactions were identified both interactively using PyMOL[27] and automatically using LIGPLOT.[28]

Table 1: Data collection and refinement statistics. Values in the parentheses correspond to the highest resolution shell

	LmrR•RBF	LmrR•ET
Data collection		
Space group	$P4_32_12$	$P2_12_12_1$
Cell dimensions		
a, b, c (Å)	35.2, 35.2, 179.7	35.7, 53.8, 144.9
α,β,γ (°)	90, 90, 90	90, 90, 90
solvent content (%)		52
Resolution range*	35 -2.4 (2.53-2.4)	50 -2.00 (2.11-2.0)
No. of unique reflections	4872	19496
Completeness (%)*	98.8 (99.3)	99.4 (100.0)
Multiplicity*	3.3 (3.4)	4.0 (4.0)
R_{merge}*	0.057 (0.37)	0.082 (0.337)
$I/\sigma I$*	15.3 (3.2)	11.5 (3.8)
Refinement		
Resolution range	30-2.35	50-2.0
R_{factor}/R_{free}	0.21/0.27	0.20/0.26
No. of atoms in asymmetric unit, average B factor (Å^2)		
Protein	809, 43.2 (chain A)	1816, 27.4 (chains A and B)
Solvent	17, 32.4	100, 13.54
Ligand	54, 49	24, 45.8
Rmsd		
Bond lengths (Å)	0.01	0.022
Bond angles (°)	1.01	1.74
Ramachandran analysis, validation		
Preferred regions (%)	100	97.3
Allowed regions (%)	0.0	2.3
Outliers (%)	0.0	0.0
Molprobity score	1.5	1.5
PDB entry	3UJY	3UJ5

$R_{merge} = \sum|I - \langle I\rangle|/|\langle I\rangle|$; $R_{factor} = \left(\sum||F_{obs}| - |F_{calc}||/|F_{obs}|\right)$, for all measured reflections in the specified resolution range; $R_{free} = \left(\sum||F_{obs}| - |F_{calc}||/|F_{obs}|\right)$, for reflections belonging to a random test set not used in refinement (10% of the data).

*Values in parentheses refer to the highest resolution shell

Drug binding assay

Dissociation constants of RBF and ET were obtained from fluorescence titration experiments. The RBF titration experiment employed the change in intrinsic fluorescence of the ligand upon binding as described.[29] The ET titration experiment employed the change in tryptophan fluorescence of LmrR upon ligand binding, following the same protocol as used previously for monitoring the binding of DAU.[8] Samples were scanned three times to obtain the final fluorescence emission spectra. Spectra were corrected for background emission and dilution effects. Apparent dissociation constants of RBF and ET were obtained by data fitting using non-linear regression analysis (using the program SigmaPlot). The RBF binding assay was performed on a Spex Fluorolog 322 fluorescence spectrophotometer at 25 °C. To a 1-ml stirred cuvette with purified LmrR in binding buffer (20 mM Tris–HCl, pH 8.0, 150 mM NaCl and 1 mM EDTA, dimer concentration of 1.5 μM), RBF was added in 2 μl steps. RBF fluorescence was measured after each titration step. For reference, RBF was also titrated to the buffer solution without protein**.** The excitation and emission wavelength were 435 and 523 nm, respectively. Tryptophan fluorescence for the ET binding assay was recorded at room temperature with an Aminco Bowman Series 2 spectrophotometer, using a 0.2 ml cuvette and an excitation wavelength of 295 nm. ET was added in 2 μl steps **to** 0.1 μM purified LmrR (dimeric concentration). Emission spectra were obtained from 300 to 450 nm and processed as above to obtain tryptophan fluorescence quenching values.

Results

Structure determination

Structures of LmrR, in complex with RBF (LmrR•RBF) and ET (LmrR•ET), were solved by molecular replacement and refined at 2.35 Å and 2.0 Å, respectively. The LmrR•RBF structure has the same crystal form (space group $P4_32_12$) as the previously determined structure of LmrR bound to H33342 (LmrR•H33342), with a single protein subunit occupying the asymmetric unit and the 2-fold axis of the LmrR dimer coinciding with a crystallographic 2-fold axis. The

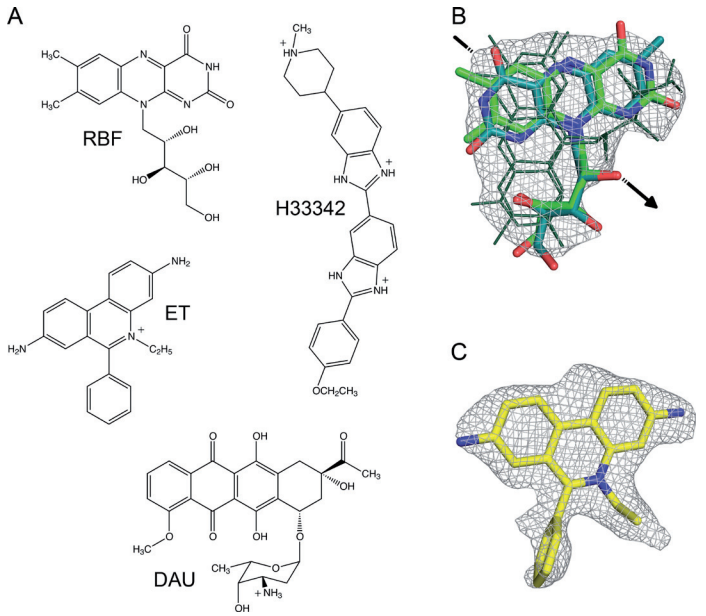

Figure 1. LmrR ligands. (A) Chemical structures of riboflavin (RBF), ethidium (ET), Hoechst 33342 (H33342) and daunomycin (DAU). (B) $2F_o$-F_c electron density for RBF in the LmrR•RBF structure calculated at 2.0 Å resolution and contoured at 1σ. The two crystallographically independent binding conformations of RBF are shown in stick representation with the carbon atoms colored green or cyan (oxygen and nitrogen atoms are colored red and blue, respectively). These binding conformations differ by a ~180° rotation of the heterocyclic isoalloxazine core relative to the ribityl moiety. The other two binding modes of RBF (shown with dark green lines) are related to the first two by 2-fold crystallographic symmetry (the location of the crystallographic dyad is indicated with an arrow). (C) $2F_o$-F_c electron density for ET in the LmrR•ET structure, calculated at 2.4 Å resolution and contoured at 1σ. The ligand is shown in stick representation with carbon, oxygen, nitrogen atoms colored yellow, red and blue, respectively.

LmrR•ET structure has a different crystal form (space group $P2_12_12_1$), with a full LmrR dimer occupying the asymmetric unit, which is identical to that of the previously determined structure of LmrR bound to DAU (LmrRŸDAU).[8] Data statistics, refinement details and model content of the two structures are summarized in Table 1. There is well-defined electron density for the polypeptide chains, except for the tip region of the β-wings (residues 70-75) and the N- and C-termini (residues 1-5 and 109-116), which show a high degree of disorder, especially in the LmrR•RBF structure. Amino acid residues with no observed electron density were excluded from the final models. The X-ray data clearly confirms the presence of the bound ligands in the central drug-

binding pore of LmrR (Figures 1 and 2). Electron density for the heterocylic aromatic core and ethyl substituent of ethidium molecule is well defined, but its phenyl substituent shows less defined density, indicative of some degree of rotational mobility (Figure 1B). Docking of RBF was more complicated since the electron density of the ligand in the LmrR•RBF structure is averaged by the occurrence of four mutually exclusive binding modes (Figure 1C). Two binding modes of RBF are related by a conformational change in the ligand, involving a ~180° flip of the heterocylic isoalloxazine moiety relative to the ribityl side chain. In addition, each of these two binding modes has a crystallographic symmetry equivalent, due to the symmetrical position of the ligand-binding site in LmrR•RBF on the crystallographic dyad. With the four binding modes superimposed, the isoalloxazine moiety of RBF shows a good fit to the averaged density, but the ribityl side chain is less well defined. Since the four binding modes are highly similar, we will describe only one in detail in the rest of this paper.

Ligand binding and recognition

The overall binding modes of ET and RBF are similar to those of H33342 and DAU in the previously determined LmrR-ligand structures (Figures 2 and 3).[8] The flat aromatic cores of the ligands slide into the centre of the flat-shaped hydrophobic pore of the LmrR dimer, while the more bulky ring substituents protrude towards the pore opening. Drug recognition by LmrR is dominated by the aromatic stacking interactions of the Trp96/Trp96' indole pair with the heterocylic aromatic cores of the ligands. Further stabilization is provided by apolar contacts of the ligands with hydrophobic amino acid residues in helices α1, α4, α1' and α4', which face the drug binding pore and surround the central Trp96/Trp96' pair. With a total pore volume of ~1200 Å^3 and ligand volumes of 300-400 Å^3, it is evident that the ligands occupy only a fraction of the volume that is available in the central pore (Figure 2C). The hydrophobic interior of the pore prohibits formation of H-bonds between protein and polar atoms in the cores of the ligands. However, while wedged inside the pore, polar atoms at the edges of the heterocyclic rings systems are accessible to solvent, and a few localized solvent molecules contribute to drug binding stabilization. Noteworthy, all protein residues in close vicinity of the ligands (within 5 Å distance) are hydrophobic (Figure 3), except in LmrR•RBF, where

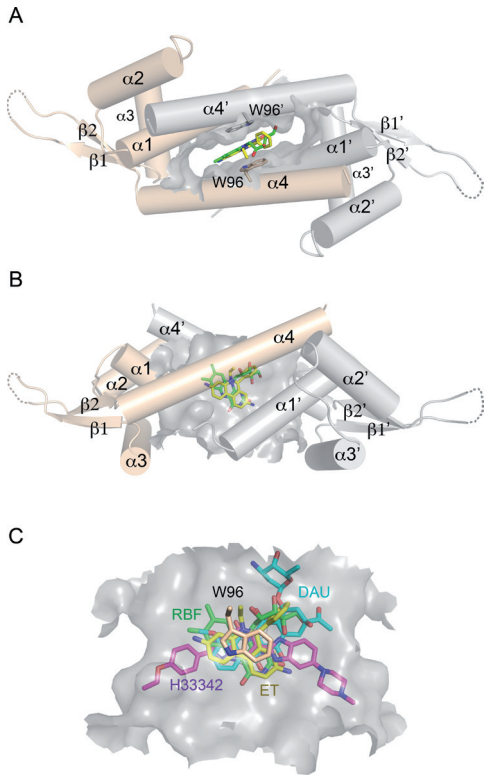

Figure 2. Overall structures of the LmrR-ligand complexes. (A) Top view of the LmrR•RBF dimer in cartoon representation, showing the central pore accessible surface (in grey) and a bound RBF (in green) stacked in between the W96/W96' residues. For clarity, only one of the four RBF binding modes is shown. An overlay of bound ET (in yellow) is added, based on a superposition of the LmrR•ET and LmrR•RBF structures. The two chains of the LmrR dimer are colored brown and grey, and the secondary structure elements are labeled. (B) Side view of the LmrR•RBF dimer, showing accessible surface for the central pore, bound RBF, and an overlay of bound ET from the LmrR•ET structure. (C) Overlays of bound RBF (carbon atoms colored green), ET (yellow), DAU (cyan) and H33342 (magenta) in the central pore of LmrR. The positions of the ligands are based on superpositions of the different ligand-bound LmrR structures. The pore surface is calculated from the LmrR•RBF structure. Oxygen and nitrogen atoms are colored red and blue, respectively.

Asp100 contributes to the stabilization of the hydrophilic ribityl moiety of the ligand, possibly by the formation of hydrogen bonds, although the observed disorder of the ribityl moiety indicates that these interactions are quite weak. A similar situation occurs in LmrR•DAU, where Asp100 is found near to the amino sugar substituent of DAU, but does not form strong directive interactions, e.g. hydrogen bonds, as indicated by the high disorder of the substituent.[8] In

LmrR•ET the carboxylate groups of Asp100 (and its dimer-related equivalent Asp100') are at ~8 Å distance from the positively charged quaternary amine of ethidium, indicating that long-range electrostatic effects may contribute to ligand binding.

Ligand binding and recognition were further analyzed by calculation of the buried accessible surface areas (ASA) and the ligand/binding site shape complementarities of the four structurally characterized LmrR-ligand complexes (Table 2).

Table 2: Ligand binding constant, ligand burial and shape complementarity in LmrR-drug complexes

Ligand	K_d (mM)	ASA ligand buried†	ASA protein buried†	%ASA ligand buried	Shape complementarity
ET	0.011 (±0.002)	380 (303/77)	324 (285/39)	73	0.73
RBF	0.25 (±0.04)	433 (236/197)	403 (323/80)	78	0.77
H33342	0.021 (±0.008)‡	565 (494/71)	476 (442/34)	78	0.68
DAU	0.24 (±0.05)‡	506 (333/173)	440 (349/91)	72	0.73

† In parenthesis are the contributions from apolar/polar atoms

‡ K_D values for H33342 and DAU are taken from Madoori *et al*.[8]

In addition, the dissociation constants (K_d) of RBF and ET were determined by fluorescence-based binding assays, and compared to those derived previously for H33342 and DAU (Figure 4 and Table 2). The results reveal similar ligand ASA burial (75 ± 3%) and ligand/binding site shape complementarity (0.73 ± 0.04) for the four different LmrR-ligand complexes. The shape complementarities are similar to those observed in ligand-bound complexes of other multidrug binding transcription regulators,[17] but the ligand ASA burial values are lower, consistent with the more open environment of the drug binding pore in LmrR. The measured K_d values (Figure 4) reveal a strong binding affinity of LmrR for the different ligands. ET shows the tightest binding (K_d = 11 ± 2 nM), comparable to that of H33342, while RBF, like DAU, shows an approximately 20-fold lower binding affinity (K_d = 0.25 ± 4 μM). The different binding affinities of ET and

H33342 versus RBF and DAU are not explained by differences in total ligand ASA burial, but there is a clear correlation between ligand binding affinity and amount of hydrophobic versus hydrophilic ASA buried by the protein upon binding the ligands. For LmrR•ET and LmrR•H33342, over 80-87 % of the total buried surface of the ligand is hydrophobic, compared to 55-66 % for LmrR•RBF and LmrR•DAU. These differences arise from the hydrophilic ribityl and amino sugar substituents in RBF and DAU, respectively, and are indicative of the penalty paid by accommodating these substituents inside the hydrophobic drug-binding pore.

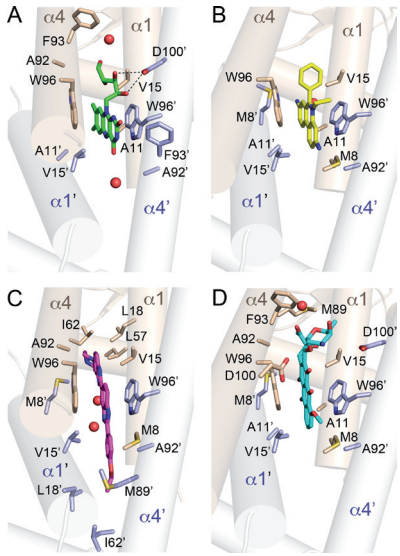

Figure 3. Close-up views of the multidrug binding site and ligand-interacting residues of (A) LmrR•RBF, (B) LmrR•ET, (C) LmrR•H33342 and (D) LmrR•DAU. Amino acid residues within a radius of 4.5 Å from a ligand are shown in stick representation and labeled. Colors as in Figure 2. Water molecules within 3.5 Å from a ligand are shown as red spheres.

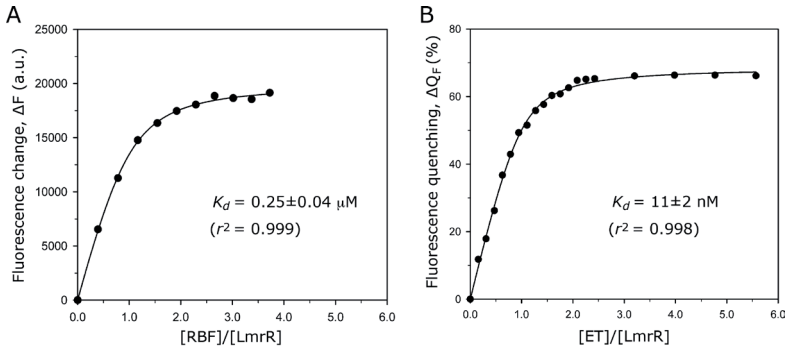

Figure 4. Fluorescence titration curves measuring RBF and ET binding to LmrR.

Conformational flexibility of LmrR

Comparison of the LmrR•RBF and LmrR•ET structures reveals minimal differences in side chain conformations at the ligand-binding site. However, as noted previously for the LmrR•H33342 and LmrR•DAU structures,[8] the LmrR•RBF and LmrR•ET dimers show significant differences in helix orientations, affecting the size and shape of the central ligand binding pore, as well as of the DNA binding site (Figures 3 and 5A). Compared to LmrR•RBF, helices α4 and α4′ in LmrR-ET have rotated outwards, away from the dimer centre, by ~10° while helices α1 and α1′ have rotated inwards by 8°. As a result the mouth of the central pore is somewhat more open in the ET-bound LmrR dimer, allowing

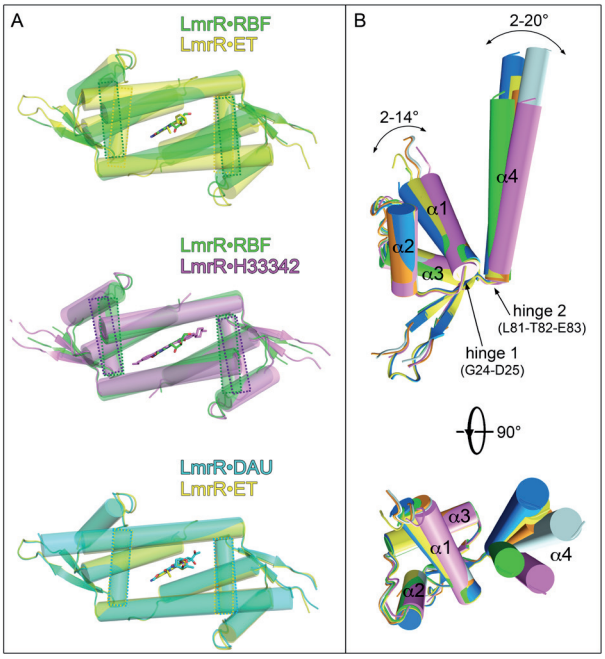

Figure 5. Conformational differences between the drug-bound structures of LmrR. (A) Pairwise superpositions of the four drug-bound dimers. The positions of the DNA recognition helices (α3, α3′) are emphasized by dashed lines. (B) Superposition of the single subunits of the four drug-bound LmrR structures in two different views, related by a 90° rotation. For the asymmetric LmrR•ET and LmrR•DAU dimers both subunits are included in the superposition. The two hinge regions are indicated by arrows and the ranges by which the orientations of helices α1 and α4 differ are shown. The coloring is as follows: subunit A and B of LmrR•ET, yellow and orange; subunit A and B of LmrR•DAU, light and dark blue, subunit A or B of LmrR•RBF, green; subunit A or B of LmrR•H33342, magenta.

a better accommodation of the ligand's phenyl group. The movement of the α1, α1′, α4 and α4′ helices is further coupled to an inward rotation of the core of the wHTH domains by ~17°, including helices α3 and α3′, shortening the distance between the centres of the DNA-binding helices from 31 Å to 27 Å. A superposition of the individual subunits of the different ligand-bound LmrR dimers reveals that the conformational flexibility largely originates from two hinge regions in the polypeptide chain of LmrR: in the α1-α2 loop and in the loop that connects α4 to the β-wing (Figure 5B). Differences in backbone conformations at the hinge regions affect the orientations of helices α1 and α4 relative to the core of the wHTH-domain. Significant differences in the relative orientations of helices α1 and α4 are also observed between the two subunits of the LmrR•ET and LmrR•DAU dimers, consistent with the presence of asymmetric dimers in the crystal with a ligand bound in a single dominant configuration.

Since LmrR-RBF and LmrR-ET crystallized in different crystal forms, it is possible that the observed conformational change is caused by differences in the crystal packing geometry. Indeed, pairwise comparison of the RBF-bound versus H33342-bound and ET-bound versus DAU-bound structures, which share the same crystal forms, reveals smaller conformational differences (Figure 5A). In particular, the overall structures of the LmrR•ET and LmrR•DAU dimers are highly similar. On the other hand, the LmrR•RBF and LmrR•H33342 dimers still reveal significant reorientations of the α-helices at the ligand and DNA binding sites, indicating that drug-specific interactions at the drug binding site of LmrR, at least to some extent, directly influence the dimeric conformational changes.

Discussion

The new LmrR•ET and LmrR•RBF structures, and their comparison with the previously determined structures of LmrR bound with H33342 and DAU, confirm that LmrR contains a single ligand-binding site within its large symmetrical flat-shaped pore, and employs a common set of almost exclusively apolar residues to achieve multidrug binding. Conformational flexibility is important for the function of LmrR as a multidrug-induced transcription regulator. The ligand-bound LmrR dimers show different conformations, affecting the overall geometries of both the drug and DNA binding sites. In particular, the overall

flexibility is associated with a change in mutual distance and inclination of the two DNA recognition helices, which likely affects their ability to simultaneously insert into two neighboring major grooves on the DNA.[8] This indicates that the observed flexibility and conformational coupling of the drug and DNA binding sites play a role in the drug-based induction mechanism of LmrR. A similar allosteric cross-talk, in which drug binding is coupled to a change in the mutual distance and inclination of the two DNA recognition helices, has been observed for other multidrug binding transcriptional regulators, although the conformational changes underlying the allosteric coupling differ significantly.[11] To which extent the changes affect the binding of LmrR to DNA cannot be validated due to lack of structural information of the LmrR•DNA complex.

The "common-site" multidrug binding mechanism of LmrR differs significantly from the canonical "multisite" model, as presented by QacR, in which multi-faceted ligand recognition is facilitated by a series of overlapping "mini-pockets" within a large drug binding site.[11] The common binding mode of the different ligands in the central pore of LmrR is largely dictated by the symmetry-related Trp96/Trp96' pair, which acts as an anchoring point and selectivity filter, due to its ability to form strong aromatic stacking interactions with the flat heterocyclic cores of the ligands. Promiscuous drug binding is realized by directing the groups of different chemistries and sizes towards the opening of the pore, where they are more solvent-exposed. The use of two subsites, one for interacting strongly with the planar and hydrophobic groups of the ligands, serving as a common anchoring point, while another, more solvent-exposed subsite accommodates the variable groups of the ligands, is reminiscent of the alternative multidrug-recognition model displayed by BmrR.[17, 18] However, the symmetrical location of the multidrug binding site within a large and flexible pore, and the unique disposition and role of the Trp96/Trp96' pair, significantly differentiate LmrR from BmrR, which uses a non-symmetrical drug binding pocket that is relatively small and rigid.

Another striking difference of LmrR with other characterized MD-linked transcription regulators like QacR and BmrR is the absence of polar or charged residues at the drug-binding site. Multidrug recognition is rarely devoid of polar contributions, even though they conflict with the necessity of a multidrug binding protein to avoid interactions displaying high specificity. MD-linked transcription regulators like QacR and BmrR use a limited set of polar residues to form complementary interactions with H-bonding or charged atoms of the

ligands, and, in particular, binding of cationic lipophilic drugs is stabilized by short-range electrostatic interactions with buried acidic residues.[11, 16, 17, 30] It should also be noted, though, that the interactions with acidic residues are not always critical for binding the cationic ligands as recently revealed for QacR.[31] In LmrR, Asp100 and its dimer-related counterpart Asp100' are too far from the drug binding site to allow short-range electrostatic interactions, although long-range electrostatic interactions could contribute to binding of the lipophilic cationic drugs, explaining the tight binding of ET and H33342. On the other hand, the similar K_d's of RBF and DAU, a neutral and positively charged molecule, respectively, indicate that electrostatic interactions with Asp100 are not a major discriminating factor in drug binding by LmrR. Rather, together with the aromatic stacking interactions of the Trp96/Trp96' pair, burial of hydrophobic surface appears to be the key component of multidrug recognition in LmrR. This is evident from the correlation between ligand binding affinity and amount of hydrophobic versus hydrophilic ASA buried by the protein upon binding the ligands. The weaker binding ligands RBF and DAU show a significantly smaller burial of hydrophobic versus hydrophilic burial than ET and H33342, due to the presence of their hydrophilic substituents. It should be noted, though, that, on average, the observed ligand binding affinities for LmrR are significantly higher than those reported for other multidrug binding transcriptional regulators, for which ligand K_d values usually fall in the 1-100 μM range. How does LmrR manage to obtain such a tight binding in the absence of polar residues at its drug-binding site? Most likely, lack of polar amino acids is compensated by the central location of the drug-binding site within a large pore, allowing easy access for solvent to the ligands from all directions within the plane of the pore. In this way, LmrR avoids a large desolvation penalty for transfer of polar groups to the hydrophobic binding site. By comparison, in multidrug binding proteins like QacR and BmrR the ligand binding pockets are more occluded from solvent, necessitating the use of polar and charged amino acid residues to avoid large destabilizing effects from burying the hydrophilic groups of the ligands.

In conclusion, our structural studies have elucidated key features of multidrug recognition in LmrR, which, in its essence, represents a unique combination of the mechanistic features of multidrug binding exhibited by other MDR-linked regulators. Like QacR, LmrR has the availability of a large and flexible region for ligand binding, but unlike QacR and like BmrR, it employs

a single set of strong aromatic interactions to anchor the structurally and chemically diverse compounds to a common ligand binding subsite, rather than different sets of hydrophobic and hydrophilic residues defining partially overlapping subsites. The nature of this common ligand binding subsite, with a symmetry-related tryptophan pair to firmly grab the flat, aromatic rings of the ligands, is a unique feature of LmrR, as is the absence of polar and charged residues in its multidrug binding site. Although difficult to quantify, water molecules must play a significant role in multidrug binding by LmrR, via their ability to form H-bonds with polar atoms of the ligands, and by decreasing the free energy of ligand binding due to the entropy gain associated with their release from the protein-drug hydrophobic interface. Whether distantly located acidic residues in LmrR are important to stabilize the cationic lipophilic ligands is still unclear. This, and the precise mechanism by which multidrug binding in LmrR affects DNA binding, are subjects of future investigations.

Acknowledgements

We thank Dirk-Jan Slotboom for his help with the LmrR-RBF fluorescence binding assay, and the staff of the ESRF (Grenoble) for providing facilities for diffraction measurements and for assistance. Work was supported in part by Ubbo Emmius Bursaries (University of Groningen) awarded to P.K.M. and A.T.

References

(1) Bolhuis, H., Molenaar, D., Poelarends, G., van Veen, H. W., Poolman, B., Driessen, A. J., and Konings, W. N. (1994) Proton motive force-driven and ATP-dependent drug extrusion systems in multidrug-resistant *Lactococcus lactis*, *J. Bacteriol. 176*, 6957-6964.

(2) Lubelski, J., de Jong, A., van Merkerk, R., Agustiandari, H., Kuipers, O. P., Kok, J., and Driessen, A. J. (2006) LmrCD is a major multidrug resistance transporter in *Lactococcus lactis*, *Mol. Microbiol. 61*, 771-781.

(3) Agustiandari, H., Lubelski, J., van den Berg van Saparoea, H. B., Kuipers, O. P., and Driessen, A. J. (2008) LmrR is a transcriptional repressor of expression of the multidrug ABC transporter LmrCD in *Lactococcus lactis*, *J. Bacteriol. 190*, 759-763.

(4) Agustiandari, H., Peeters, E., de Wit, J. G., Charlier, D., and Driessen, A. J. (2011) LmrR-mediated gene regulation of multidrug resistance in *Lactococcus lactis*, *Microbiology 157*, 1519-1530.

(5) De Silva, R. S., Kovacikova, G., Lin, W., Taylor, R. K., Skorupski, K., and Kull, F. J. (2005) Crystal structure of the virulence gene activator AphA from *Vibrio cholerae* reveals it is

a novel member of the winged helix transcription factor superfamily, *J. Biol Chem. 280*, 13779-13783.

(6) Gury, J., Barthelmebs, L., Tran, N. P., Diviès, C., and Cavin, J.-F. (2004) Cloning, deletion, and characterization of PadR, the transcriptional repressor of the phenolic acid decarboxylase-encoding *padA* gene of *Lactobacillus plantarum*, *Appl. Environ. Microbiol. 70*, 2146-2153.

(7) Huillet, E., Velge, P., Vallaeys, T., and Pardon, P. (2006) LadR, a new PadR-related transcriptional regulator from *Listeria monocytogenes*, negatively regulates the expression of the multidrug efflux pump MdrL, *FEMS Microb. Lett. 254*, 87-94.

(8) Madoori, P. K., Agustiandari, H., Driessen, A. J. M., and Thunnissen, A. M. W. H. (2009) Structure of the transcriptional regulator LmrR and its mechanism of multidrug recognition., *EMBO J. 28*, 156-166.

(9) Alekshun, M. N., Levy, S. B., Mealy, T. R., Seaton, B. A., and Head, J. F. (2001) The crystal structure of MarR, a regulator of multiple antibiotic resistance, at 2.3 Å resolution, *Nat. Struct. Biol. 8*, 710-714.

(10) Ellison, D. W., and Miller, V. L. (2006) Regulation of virulence by members of the MarR/SlyA family, *Curr. Opin. Microbiol. 9*, 153-159.

(11) Schumacher, M. A., Miller, M. C., Grkovic, S., Brown, M. H., Skurray, R. A., and Brennan, R. G. (2001) Structural mechanisms of QacR induction and multidrug recognition, *Science 294*, 2158-2163.

(12) Schumacher, M. A., and Brennan, R. G. (2003) Deciphering the molecular basis of multidrug recognition: crystal structures of the *Staphylococcus aureus* multidrug binding transcription regulator QacR, *Research Microbiol. 154*, 69-77.

(13) Schumacher, M. A., Miller, M. C., and Brennan, R. G. (2004) Structural mechanism of the simultaneous binding of two drugs to a multidrug binding protein, *EMBO J. 23*, 2923-2930.

(14) Alguel, Y., Meng, C., Teran, W., Krell, T., Ramos, J. L., Gallegos, M. T., and Zhang, X. (2007) Crystal structures of multidrug binding protein TtgR in complex with antibiotics and plant antimicrobials, *J. Mol. Biol. 369*, 829-840.

(15) Zheleznova, E. E., Markham, P. N., Neyfakh, A. A., and Brennan, R. G. (1999) Structural basis of multidrug recognition by BmrR, a transcription activator of a multidrug transporter., *Cell 96*, 353-362.

(16) Newberry, K. J., Huffman, J. L., Miller, M. C., Vazquez-Laslop, N., Neyfakh, A. A., and Brennan, R. G. (2008) Structures of BmrR-drug complexes reveal a rigid multidrug binding pocket and transcription activation through tyrosine expulsion, *J. Biol. Chem. 283*, 26795-26804.

(17) Bachas, S., Eginton, C., Gunio, D., and Wade, H. (2011) Structural contributions to multidrug recognition in the multidrug resistance (MDR) gene regulator, BmrR, *Proc. Natl. Acad. Sci. U.S.A. 108*, 11046-11051.

(18) Wade, H. (2010) MD recognition by MDR gene regulators., *Curr. Op. Struct. Biol. 20*, 489-496.

(19) Leslie, A. G. (2006) The integration of macromolecular diffraction data, *Acta Crystallogr. D Biol. Crystallogr. 62*, 48-57.

(20) Kabsch, W. (2010) Integration, scaling, space-group assignment and post-refinement, *Acta Crystallogr. D Biol. Crystallogr. 66*, 133-144.

(21) Adams, P. D., Afonine, P. V., Bunkoczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L. W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C., and Zwart, P. H. (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution, *Acta Crystallogr. D Biol. Crystallogr. 66*, 213-221.

(22) Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) Refinement of macromolecular structures by the maximum-likelihood method, *Acta Crystallogr. D Biol. Crystallogr*. *53*, 240-255.

(23) Emsley, P., and Cowtan, K. (2004) Coot: model-building tools for molecular graphics, *Acta Crystallogr. D Biol. Crystallogr. 60*, 2126-2132.

(24) Painter, J., and Merritt, E. A. (2006) TLSMD web server for the generation of multi-group TLS models, *J. Appl. Crystallogr. 39*, 109-111.

(25) Chen, V. B., Arendall, W. B., 3rd, Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S., and Richardson, D. C. (2010) MolProbity: all-atom structure validation for macromolecular crystallography, *Acta Crystallogr. D Biol. Crystallogr. 66*, 12-21.

(26) Binkowski, T. A., Naghibzadeh, S., and Liang, J. (2003) CASTp: Computed Atlas of Surface Topography of proteins, *Nucleic Acids Res. 31*, 3352-3355.

(27) DeLano, W. L. (2002) The Pymol Molecular Graphics System, DeLano Scientific, San Carlos, CA.

(28) Laskowski, R. A., and Swindells, M. B. (2011) LigPlot+: Multiple Ligand-Protein Interaction Diagrams for Drug Discovery, *J. Chem. Inf. Model*. *51* (10), 2778–2786

(29) Duurkens, R. H., Tol, M. B., Geertsma, E. R., Permentier, H. P., and Slotboom, D. J. (2007) Flavin binding to the high affinity riboflavin transporter RibU, *J. Biol. Chem. 282*, 10380-10386.

(30) Grkovic, S., Hardie, K. M., Brown, M. H., and Skurray, R. A. (2003) Interactions of the QacR multidrug binding protein with structurally diverse ligands: implications for the evolution of the binding pocket, *Biochemistry 42*, 15226-15236.

(31). Peters, K. M., Brooks, B. E., Schumacher, M. A., Skurray, R. A., Brennan, R. G. & Brown, M. H. (2011). A single acidic residue can guide binding site selection but does not govern QacR cationic-drug affinity. *PloS one* **6**, e15974.

(32) Winn MD, Ballard CC, Cowtan KD, Dodson EJ, Emsley P, et al. (2011) Overview of the CCP4 suite and current developments. Acta Crystallogr D Biol Crystallogr 67: 235-242

Chapter

Purification, crystallization and preliminary X-ray diffraction analysis of the lytic transglycosylase MltF from *Escherichia coli*

4

Pramod K. Madoori and Andy-Mark W. H. Thunnissen

Published in: *Acta crystallographica F* (2010). F66, 534-538

Abstract

The lytic transglycosylase MltF from *Escherichia coli* is an outer membrane-bound, periplasmic protein with two domains, a C-terminal catalytic domain with a lysozyme-like fold and an N-terminal domain of unknown function that is homologous to periplasmic substrate-binding proteins of ABC transporters. To investigate its structure and function, a soluble form of full length MltF (sMltF), containing both domains, and a soluble fragment containing only its N-terminal domain (sMltF-NTD), were purified and crystallized. Crystals of sMltF belonged to space group $P4_32_12$ or $P4_12_12$, with unit-cell dimensions a = b = 110.8 Å, c = 163.5 Å and two molecules per asymmetric unit. A complete data set was collected to 3.5 Å resolution. Crystals of sMltF-NTD belonged to space group $P3_121$ with unit-cell parameters a = b = 82.4 Å, c = 75.2 Å and one molecule per asymmetric unit. For sMltF-NTD a complete native data set was collected to 2.20 Å resolution. In addition, for phasing purposes a three-wavelength MAD data set was collected to 2.5 Å resolution using a bromide-soaked sMltF-NTD crystal. Using phases derived from the Br-MAD data it was possible to build a partial model of sMltF-NTD.

1. Introduction

The viability and shape of bacteria depends on the presence of an intact cell wall that surrounds their cytoplasmic membrane. The integral component of the bacterial cell wall is a heteropolymer known as peptidoglycan (PG) or murein. It is composed of glycan strands of alternatingly β-1,4-linked N-acetylmuramic acid (MurNAc) and N-acetylglucosamine (GlcNAc) residues, cross-linked by peptides that are connected to the lactyl groups of the MurNAc residues (Vollmer *et al.*, 2008). The mesh-like PG structure gives the cell wall its mechanical strength, allowing bacterial cells to withstand the high internal osmotic pressure. Once synthesized, however, the PG polymer is not a static macromolecule, but it is subject to continuous remodeling and turnover (Park & Uehara, 2008). In particular, PG cleavage is required to create space for the insertion of new material and to recycle old material during cell growth, to incise the cell wall during cell division, and to create local openings in the cell wall for allowing the insertion of various cell envelop spanning structures (Holtje, 1998, Koraimann, 2003, Scheurwater *et al.*, 2008). PG cleavage is carried out by bacterial glycolytic and peptidolytic enzymes that are referred to as autolysins. Some of these bacterial enzymes are crucial for bacterial pathogenicity and have been shown to modulate muropeptide release and/or host innate immune responses (Lee *et al.*, 2009).

Lytic transglycosylases (LTs) form one set of autolysins, that target the β-1,4-linkages between the MurNAc and GlcNAc residues of PG (Höltje, 1996, Scheurwater *et al.*, 2008). They act like lysozymes, and other β-1,4-glycosyl hydrolases, but differ with respect to the reaction products. Strictly speaking, LTs are glycosyl transferases, and not hydrolases, which combine cleavage of an inter-residue β-1,4-glycosidic bond with formation of an intra-residue 1,6-glycosidic bond, thereby producing GlcNAc-anhydro-muropeptides (Figure 1). LTs are ubiquitous among all eubacteria that produce PG, but the complement of enzymes produced by *Escherichia coli* has been the most extensively examined. *E. coli* is known to produce six outer membrane-bound lytic transglycosylases (MltA, MltB, MltC, MltD, MltE, MltF) and one soluble lytic transglycosylase (Slt70) (for reviews see (Höltje, 1996, Scheurwater *et al.*, 2008)). Most appear to act as exo-enzymes releasing GlcNAc-anhydro-muropeptides from the ends of glycan strands, except for MltE, which has been shown to be endo-acting (Kraft *et al.*, 1998). Collectively, these enzymes form

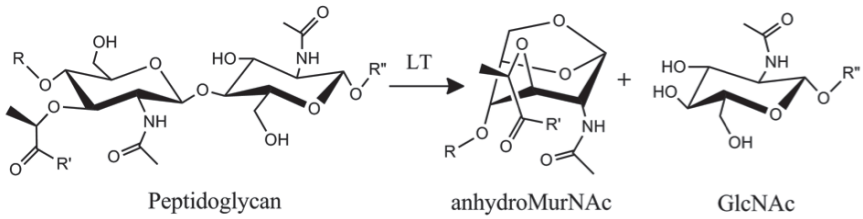

Figure 1. LTs catalyze the cleavage of the β-1,4-glycosidic bond between MurNAc and GlcNAc residues in peptidoglycan with the concomitant formation of a 1,6-anhydro-bond in the MurNAc residue.

the archetypes for three of the four families of LTs identified by Blackburn and Clarke (Blackburn & Clarke, 2001). Crystal structures, complemented with sugar and muropeptide binding studies, have been reported for Slt70 (Thunnissen *et al.*, 1994, van Asselt, Thunnissen *et al.*, 1999), MltA (van Straaten *et al.*, 2007, van Straaten *et al.*, 2005) and Slt35 (a soluble proteolytic fragment of MltB (van Asselt, Dijkstra *et al.*, 1999, van Asselt *et al.*, 2000), representing LT families I, II and III, respectively. These crystallographic studies allowed a thorough understanding of the structures and catalytic mechanism of LTs, and revealed that most of these enzymes, with the exception of MltA, share a catalytic domain that resembles the fold of goose-type lysozyme (Thunnissen *et al.*, 1995). However, the specific roles of the different *E. coli* LTs in PG metabolism has remained unclear (Heidrich *et al.*, 2002), which is emphasized by the notion that most LTs contain additional, non-catalytic domains for which the function is often unknown.

MltF from *E. coli* is a recently characterized member of LT family I, which, based on sequence analysis and functional assays, contains a typical lysozyme-like C-terminal domain (hereafter named the LT domain) that is responsible for the LT activity (Scheurwater & Clarke, 2008). As a unique feature, however, it contains an N-terminal domain that is homologous to periplasmic substrate-binding proteins of ABC transporters, in particular to those specific for histidine, lysine-arginine-ornithine (LAO) and glutamine (Tam & Saier, 1993). The function of this N-terminal domain (MltF-NTD) is unknown. No peptidoglycan binding activity could be measured for MltF-NTD, nor have any ligands been identified that may form substrates of this domain (Scheurwater & Clarke, 2008). The N-terminal domain was shown to modulate the lytic activity of the LT domain, to permit continued lysis of insoluble peptidoglycan at a

constant rate (Scheurwater & Clarke, 2008), but how this modulation happens is currently not understood.

To obtain insights in the role of the N-terminal domain of MltF, and how it may affect the catalytic function of the LT domain, we study MltF using X-ray crystallographic and biochemical methods. In this paper we describe the purification, crystallization and preliminary X-ray analysis of two soluble, C-terminally His_6-tagged forms of MltF lacking residues 1-22, one containing both domains (sMltF) and one containing only the N-terminal domain (sMltF-NTD).

2. Materials and methods

2.1 Expression and purification

Soluble MltF (sMltF, 511 residues), lacking the predicted signal sequence and transmembrane helix (residues 1-22) but with an extra C-terminal H_6-tag, was expressed using the previously published expression vector pACES-8 (Scheurwater & Clarke, 2008). Expression was carried out in the *E. coli* strain Rosetta 2 (DE3) pLysS (Novagen). A 2 liter LB culture, supplemented with chloramphenicol (34 µg ml^{-1}) and kanamycin (50 µg ml^{-1}) was incubated at 310 K until the OD_{600nm} reached ~0.6. The cells were then induced by the addition of 1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) and incubated for an additional 3 hours at 310 K. For the preparation of soluble fractions, cultured cells were harvested by centrifugation at 6000*g* for 20 min at 277 K and the resulting bacterial pellet was resuspended in 50 ml ice-cold lysis buffer containing 20 mM Tris–HCl, pH 8.0, 300 mM NaCl, 2 mM imidazole, 0.2% NP40, 1 mM β-mercaptoethanol, and appropriate amounts of DNase, RNase and protease inhibitors (Roche Applied Science). Cells were lysed using a French press, and the soluble proteins were collected by centrifugation at 6000*g* for 20 min at 277 K. The supernatant was applied to a 0.5 ml Ni-NTA (Qiagen) column pre-equilibrated with 20 mM Tris-HCl pH 8.0, 300 mM NaCl, 10 mM imidazole and 1 mM β-mercaptoethanol (buffer A). The column was washed with 3-4 column volumes of buffer A to remove unbound proteins and sMltF was eluted with 200 mM imidazole in buffer A. Elution fractions containing sMltF were pooled, diluted 6 fold in buffer B (20 mM Tris-HCl, pH 8, 1 mM EDTA and 1

mM dithiothreitol (DTT)) and subsequently loaded onto a monoQ column (GE Healthcare), which was equilibrated with buffer B. Elution was carried out with a gradient of NaCl concentrations increasing from 50 mM to 500 mM. The peak fractions containing sMltF were pooled and concentrated to 12 mg ml^{-1} in 20 mM Tris-HCl, pH 8.0, 50 mM NaCl, 1 mM EDTA and 1 mM DTT, using an Amicon ultrafiltration centrifugal device (Millipore).

Expression and purification of the N-terminal domain of sMltF (sMltF-NTD, residues 23-270 with additional C-terminal H_6-tag) followed a similar procedure as used for the full length protein. Expression was carried out with the vector pACES-13 (Scheurwater & Clarke, 2008) in C43 (DE3) *E. coli* cells, using LB media supplemented with kanamycin. A three-step purification protocol, using Ni-NTA, monoQ and gel filtration chromatography, was applied to obtain pure protein. The Ni-NTA and monoQ purification steps were performed as for sMltF. Gel filtration was carried out on a Superdex 200 column (GE Healthcare) pre-equilibrated with column buffer containing 50 mM Tris-HCl, pH 8, 50 mM NaCl and 1 mM DTT. The peak fractions containing sMltF-NTD were pooled and concentrated to 6 mg ml^{-1} in gel filtration column buffer. Protein concentrations were estimated from the absorbance at 280nm (A_{280}) using theoretical molar extinction coefficients of 84230 M^{-1} cm^{-1} and 38390 M^{-1} cm^{-1} for sMltF and for sMltF-NTD, respectively.

All purification steps were performed at 280 K and the results of each step were monitored by SDS–PAGE. The finally obtained protein samples were highly pure (>98%) and mono-disperse, as judged from silver-stained SDS-PAGE gels and dynamic light scattering experiments (DynaPro, Wyatt Technology), respectively. After concentration the protein samples were quickly frozen in liquid nitrogen or used immediately for crystallization screening.

2.2 Crystallization

Screening for initial crystallization conditions was performed using the sitting-drop vapour-diffusion method, with the aid of an Oryx-6 crystallization robot (Douglas Instruments) at 298K and using the commercial JCSG+ and PACT crystallization screens (Molecular Dimensions Ltd). Lead conditions for crystallization were further optimized by changing salt concentration, precipitant concentration, temperature, and by changing the buffering agents. The crystal optimization experiments were performed manually using the

hanging-drop vapour-diffusion method, by mixing and equilibrating equal volumes (1 µl) of the protein and reservoir solutions against 500 µl reservoir solutions in a 24-well plate. Tetragonal crystals of sMltF measuring 80 × 40 × 20 μm^3 were grown from 0.1 M NH_4-acetate, 0.1 M bis-Tris, pH 5.5, 15 % PEG 10000. Trigonal crystals of sMltF-NTD having dimensions of 200 × 60 × 60 μm^3 were grown from 0.15 M Li_2SO_4, 0.1 M Na citrate, pH 5.5, 20% PEG 3350.

2.3 X-ray data collection and processing

X-ray diffraction data were collected at the ESRF, Grenoble, using cryo-cooled crystals. Full-length sMltF crystals were cryo-protected by increasing the PEG 10000 concentration to 30 %. Cryo-protection of the sMltF-NTD crystals required the addition of 15 % glycerol to the crystallization solution. Data were integrated using *XDS* (Kabsch, 1993), and scaled and merged to unique data sets with the programs *SCALA* and *TRUNCATE* from the CCP4 suite (Win, *et al*., 2011). The sMltF crystals suffered from extensive radiation damage, resulting in a somewhat poor overall quality of the data set and a useful resolution of only 3.5 Å, even though diffraction extended to about 2.5 Å at the beginning of the data collection experiment. The sMltF-NTD crystals, on the other hand, were very stable in the X-ray beam and diffracted to 2.2 Å resolution. In addition to a native data set, a three-wavelength Br-MAD data set was collected from a single MltF-NTD crystal, that was soaked for 15-20 seconds in a solution containing 20% glycerol and 0.6 M NaBr just prior to freezing, following published protocols (Dauter *et al*., 2000) . Tables 1 and 2 list the relevant data collection statistics.

3. Results and discussion

Both full-length sMltF and sMltF-NTD were successfully purified and crystallized. X-ray data were collected on cryo-cooled crystals using the MX beam lines at the ESRF, Grenoble. Crystals of sMltF diffracted to a maximum resolution of 2.5 Å, but due to radiation damage the finally obtained unique data set was complete only up to 3.5 Å resolution (Table 1). The space group was identified as $P4_32_12$ or $P4_12_12$, with unit cell dimensions a = b = 110.8 Å, c = 163.5 Å. Computation of the Matthews coefficient indicated that the

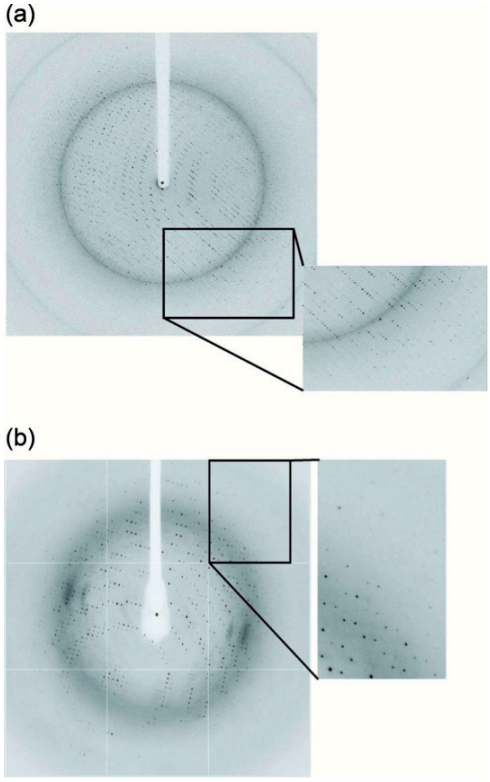

Figure 2. Diffraction images of sMltF crystals from (a) full-length protein and (b) sMltF-NTD. The edge of the detector corresponds to 1.8 Å and 1.7 Å for (a) and (b), respectively. Insets display the quality and maximum resolution of the diffraction.

asymmetric unit contains either one protein molecule (Matthews coefficient of 4.4 $Å^3$ Da^{-1}) or two protein molecules (Matthews coefficient of 2.2 $Å^3$ Da^{-1}), with a solvent content of 72% or 44% respectively. A Patterson self-rotation map did not reveal the presence of any rotational noncrystallographic symmetry (NCS), nor was any translational NCS detected in a native Patterson map, indicating that the asymmetric unit probably contains a single protein molecule. In solution, sMltF behaves as a monomer based on gel-filtration chromatography and static light-scattering analysis (not shown). It cannot be excluded, however that an NCS peak that is present in the self-rotation map is obscured by a crystallographic symmetry-axis peak. Crystals of sMltF-NTD allowed the collection of a complete data set to 2.2 Å resolution (Table 2; Fig. 2b).

Based on these data, the space group of the sMltF-NTD crystals was initially determined to be $P3_121$ or $P3_221$, with unit cell dimensions a = b = 82.4 Å, c = 75.2 Å and a single molecule per asymmetric unit (solvent content 51 %).

Molecular replacement was tried as a method to obtain initial phases for the sMltF and sMltF-NTD diffraction data, using search models based on the LT domain of Slt70 and on various structures of periplasmic substrate-binding proteins, but without success. However, in an alternative approach to obtain phases, a three-wavelength MAD data set was collected to 2.5 Å resolution of a single bromide-soaked crystal of sMltF-NTD (Table 2). Phase calculation and refinement were performed for both space groups ($P3_121$ or $P3_221$) using the program *SHARP/autoSHARP* (Vonrhein *et al.*, 2007), followed by density modification with *SOLOMON* (Abrahams & Leslie, 1996). Three different bromide sites were identified in the asymmetric unit and the best set of phases was calculated using the space group $P3_121$. The overall figure-of-merit (FOM) was 0.43 and 0.90 before and after solvent flipping, respectively, for reflections in the resolution range between 71.7 Å and 3.0 Å. The resulting experimental electron density map showed clear solvent-protein boundaries and in the

Table 1. Summary of the X-ray data collection for full-length sMltF. Values in parentheses are for the highest resolution shell.

Beam line	ID23-2
Space group	$P4_32_12$ or $P4_12_12$
Wavelength (Å) Unit cell parameters (Å)	**0.873**
a=b	110.8
c	163.5
Solvent content (%)	44.4
Resolution range (Å)	49.5-3.5
Total number of observations	55062 (7923)
No. unique reflections	12980 (1864)
Multiplicity	4.2 (4.3)
Completeness (%)	97.2 (98.0)
R_{merge} [§] (%)	12.3 (21.5)
Mean $I/s(I)$	8.6 (5.9)

[§] $R_{merge} = \sum_{hkl} \sum_i |I_i(hkl) - \langle I(hkl) \rangle| \; / \; \sum_{hkl} \sum_i I_i(hkl)$, where $I_i(hkl)$ is the *i*th observation of reflection *hkl* and $\langle I(hkl) \rangle$ is the weighted average intensity for all observations *i* of reflection hkl.

protein-associated densities the features of secondary structural elements were clearly visible. Using automated-model building (Terwilliger, 2003) it was possible to fit a partial model of nearly 137 amino-acids (52 % of the complete protein) in the electron map. Further model building and refinement of the sMltF-NTD structure is in progress and will be reported elsewhere. In addition, crystallization conditions of full-length sMltF are currently being optimized in order to obtain better quality crystals.

Table 2. Summary of the X-ray data collection for sMltF-NTD. Values in parentheses are for the highest resolution shell.

	Native	Br-MAD		
		peak	inflection	remote
Beam line	ID29	ID29		
Space group	$P3_121$	$P3_121$		
Unit cell parameters (Å)				
a=b	82.4	82.6		
c	75.2	75.2		
Wavelength (Å)	0.9300	0.9198	0.9206	0.8569
Resolution range (Å)	71.4-2.2	41.4-2.6	41.3-2.7	41.4-2.8
Unique reflections	15240	9004	8503	7526
Multiplicity	7.0	8.0	8.0	8.0
Completeness (%)	100 (98.3)	100 (98.5)	100 (99)	100 (99)
R_{merge} [§] (%)	5.5 (47.0)	8.0 (48.1)	10.6 (64.0)	9.5 (53.0)
Mean *I*/*s*(*I*)	19.1 (4.2)	20.5 (4.6)	17.0 (3.5)	18.2 (4.2)

[§] $R_{merge} = \Sigma_{hkl} \Sigma_i |I_i(hkl) - \langle I(hkl) \rangle| / \Sigma_{hkl} \Sigma_i I_i(hkl)$, where $I_i(hkl)$ is the *i*th observation of reflection *hkl* and $\langle I(hkl) \rangle$ is the the weighted average intensity for all observations *i* of reflection hkl.

Acknowledgements

The authors thank Edie M. Scheurwater and Anthony J. Clarke (University of Guelph, Canada) for providing the expression vectors for producing the soluble MltF proteins. This research was supported in part by an Ubbo Emmius Bursary, University of Groningen, awarded to the first author.

References

Abrahams, J. P. & Leslie, A. G. (1996). Acta Crystallogr D Biol Crystallogr 52, 30-42.
Blackburn, N. T. & Clarke, A. J. (2001). J Mol Evol 52, 78-84.
Dauter, Z., Dauter, M. & Rajashankar, K. R. (2000). Acta Crystallogr D Biol Crystallogr 56, 232-237.
Heidrich, C., Ursinus, A., Berger, J., Schwarz, H. & Höltje, J.-V. (2002). J Bacteriol 184, 6093-6099.
Holtje, J. V. (1998). Microbiol Mol Biol Rev 62, 181-203.
Höltje, J. V. (1996). EXS 75, 425-429.
Kabsch, W. (1993). J Appl Crystallogr 26, 795-800.
Koraimann, G. (2003). Cellular and Molecular Life Sciences (CMLS) 60, 2371-2388.
Kraft, A. R., Templin, M. F. & Holtje, J. V. (1998). J Bacteriol 180, 3441-3447.
Lee, M., Zhang, W., Hesek, D., Noll, B. C., Boggess, B. & Mobashery, S. (2009). J Am Chem Soc 131, 8742-8743.
Park, J. T. & Uehara, T. (2008). Microbiol Mol Biol Rev 72, 211-227, table of contents.
Scheurwater, E., Reid, C. W. & Clarke, A. J. (2008). Int J Biochem Cell Biol 40, 586-591.
Scheurwater, E. M. & Clarke, A. J. (2008). J Biol Chem 283, 8363-8373.
Tam, R. & Saier, M. H., Jr. (1993). Microbiol Rev 57, 320-346.
Terwilliger, T. C. (2003). Acta Crystallogr D Biol Crystallogr 59, 1174-1182.
Thunnissen, A. M., Dijkstra, A. J., Kalk, K. H., Rozeboom, H. J., Engel, H., Keck, W. & Dijkstra, B. W. (1994). Nature 367, 750-753.
Thunnissen, A. M., Isaacs, N. W. & Dijkstra, B. W. (1995). Proteins 22, 245-258.
van Asselt, E. J., Dijkstra, A. J., Kalk, K. H., Takacs, B., Keck, W. & Dijkstra, B. W. (1999). Structure 7, 1167-1180.
van Asselt, E. J., Kalk, K. H. & Dijkstra, B. W. (2000). Biochemistry 39, 1924-1934.
van Asselt, E. J., Thunnissen, A. M. & Dijkstra, B. W. (1999). J Mol Biol 291, 877-898.
van Straaten, K. E., Barends, T. R., Dijkstra, B. W. & Thunnissen, A. M. (2007). J Biol Chem 282, 21197-21205.
van Straaten, K. E., Dijkstra, B. W., Vollmer, W. & Thunnissen, A. M. W. H. (2005). Journal of Molecular Biology 352, 1068-1080.
Vollmer, W., Blanot, D. & de Pedro, M. A. (2008). FEMS Microbiology Reviews 32, 149-167.
Vonrhein, C., Blanc, E., Roversi, P. & Bricogne, G. (2007). Methods Mol Biol 364, 215-230.
Winn, M. D., Ballard, C. C., Cowtan, K. D., Dodson, E. J., Emsley, P., Evans, P. R., Keegan, R. M., Krissinel, E. B., Leslie, A. G., McCoy, A., McNicholas, S. J., Murshudov, G. N., Pannu, N. S., Potterton, E. A., Powell, H. R., Read, R. J., Vagin, A. & Wilson, K. S. (2011). Acta Crystallogr D Biol Crystallogr 67, 235-242

Chapter

Crystal structure of the SBP-like N-terminal domain of *Escherichia coli* lytic transglycosylase MltF

5

Pramod Kumar Madoori, Edie M. Scheurwater, Anthony J. Clarke and Andy-Mark W. H. Thunnissen

Manuscript in preparation

ABSTRACT

The lytic transglycosylase MltF from *Escherichia coli* contains, in addition to a C-terminal lysozyme-like catalytic domain, an N-terminal domain that is homologous to periplasmic substrate-binding proteins (SBPs) of ABC transporters. Neither the functional role nor the ligand binding characteristics of the SBP-like N-terminal domain are known. Here we report the structure of the N-terminal domain of MltF (MltF-NTD) solved by X-ray crystallography at 2.2 Å resolution. Its bilobal architecture, the topology of the central β-sheets in its two subdomains and the structure of its inter-subdomain linkers define MltF-NTD as a type-2 SBP. The domain is particularly similar to periplasmic SBPs that bind polar and charged amino acids, although the sequence identity is low (15-24%). The structure has an open conformation and no bound ligand present at the inter-subdomain cleft. A comparison with the "ligand-closed" structures of the glutamine-binding, the lysine/arginine/ornithine-binding and histidine-binding proteins revealed that two of the residues that bind the α-ammonium and α-carboxylate groups of the ligand in the amino-acid-binding proteins are conserved in MltF-NTD (Arg125 and Asp203). However, there are also important differences in side chain composition and local polypeptide conformations at the inter-subdomain cleft, indicating that MltF-NTD has significantly different ligand binding characteristics as its structural homologs.

INTRODUCTION

Substrate-binding proteins (SBPs) form a wide-spread superfamily of proteins and domains, which are found in all three kingdoms of life and function as receptors in signal transduction, regulation and solute transport.[1-7] Despite their low overall sequence identities and large variation in sizes, SBPs share a highly similar overall fold, consisting of two structurally conserved α/β domains, that are connected by a hinge region and that form a deep cleft at their interface for substrate binding.[6] Hinge-bending movements lead to closure of the substrate-binding site, thereby burying a bound substrate at the domain interface. This mode of substrate binding has been referred to as the "venus fly trap mechanism".[8] SBPs can be divided into various groups using different classification schemes. Early classifications are mainly based on amino acid sequence information and knowledge of substrate specificity (i.e., metal-, carbohydrate-, vitamin-, tetrahedral oxyanions-, compatible solutes-, amino acids- and peptide-binding proteins).[7] Alternatively, SBPs of known three-dimensional structure are commonly divided into three types, based on the central β-sheet topology of the two α/β domains and the architecture of the hinge region.[9] A new structure-based classification, recently proposed by Berntsson et al.,[10] further divides the three types into eight clusters, providing a more precise definition of the structural relationships among the different SBPs.

In Gram-negative bacteria the majority of SBPs are located as soluble proteins in the periplasmic space (hence, this group of SBPs are also named periplasmic binding proteins), where they mainly associate with ATP-binding cassette (ABC)-transporters. As such they serve a central role in substrate translocation by binding their ligands selectively and delivering them to the translocation machinery in the membrane.[2] Surprisingly, an SBP-like domain has also been reported to occur in a periplasmic protein not associated with solute transport, i.e. in the lytic transglycosylase family member MltF.[11] Lytic transglycosylases (LTs) are specialized peptidoglycan-cleaving enzymes that have a role in cell wall remodeling and recycling during bacterial growth and cell division.[12] They are ubiquitous among all eubacteria that produce PG, and are often present in several distinct forms within the same species.[12-15] For *Escherichia coli*, MltF is the seventh LT family member that has been functionally characterized, and genomic sequence analysis show that it is also present in various other eubacteria. Like most LTs, MltF is an outer-membrane-attached

periplasmic enzyme that contains a typical lysozyme-like C-terminal domain responsible for its peptidoglycan degrading activity.[16] Unlike all other LTs, however, it also contains an N-terminal domain homologous to the periplasmic SBPs of ABC transporters, in particular to those specific for histidine, lysine/arginine/ornithine and glutamine. The function of this N-terminal domain (MltF-NTD) is unknown. [11]

To reveal its precise structural relationships with the SBPs and to find clues for unraveling its biological function, we set out to determine the three-dimensional structure of MltF-NTD using X-ray crystallography. This study reports the crystal structure of MltF-NTD at 2.2 Å resolution.

EXPERIMENTAL PROCEDURES

Structure determination

Expression, purification and crystallization of the soluble N-terminal domain of *E. coli* MltF (UniProt accession number P0AGC5), without the N-terminal signal sequence and carrying a C-terminal poly-histidine tag) has been published elsewhere.[17] Initially, a partial structure of MltF-NTD was built in electron density maps at 3.0 Å resolution, obtained from Br-MAD diffraction experiments as described.[17] This partial structure, which was only 52% complete, was then improved and extended using native diffraction data collected at 2.2 Å resolution. First, molecular replacement was carried out with the program *Phaser*.[18] After obtaining a clear solution for the orientation and position of the partial structure in the asymmetric unit, the model was extended by automated model building using the program *Resolve*[19] with data from 25 to 2.2 Å resolution. Ninety-three percent of the complete model was built automatically, and the remainder was added iteratively using manual building with the program *Coot*,[20] alternated by restrained-model refinement using the programs *Refmac5*,[21] and, in the final stages, *Phenix.refine*.[22] For cross-validation purposes using R_{free} a randomly selected subset of reflections (5%) was left out from the refinements. TLS refinement was included in the final refinement runs.[23] Water molecules were added using Coot based on stringent electron density and interaction geometry criteria. The final structure was refined to an R_{work} and R_{free} of 19.5% and 23.4%, respectively. Model assessment was carried out with Molprobity.[24] Table 1 summarizes the refinement statistics.

Sequence and structure analysis

Close homologs of *E. coli* MltF and multiple sequence alignments for sequence conservation analysis were obtained using BLAST, the JPRED server (http://www.compbio.dundee.ac.uk/www-jpred/)[25] and the Consurf server (http://consurf.tau.ac.il/).[26]

RESULTS AND DISCUSSION

Structure determination

The structure of MltF-NTD (residues 23–270 of *E. coli* MltF, lacking the N-terminal signal peptide and including a C-terminal poly-histidine tag of sequence KLAAALEHHHHHH) was solved at 2.2 Å resolution by the multiple wavelength anomalous dispersion (MAD) technique using a bromide-soaked sMltF-NTD crystal. The crystallization procedure and diffraction data statistics were published elsewhere.[17] Refinement details and model content are summarized in Table 1. The structure of MltF-NTD is well defined, except for the first eleven residues at the N-terminus (residues 23-33, sequence numbering is according to full length *E. coli* MltF with signal peptide) and the last seven residues of the poly-histidine tag. These regions appeared to be disordered and were therefore not included in the model.

Overall structure of MltF-NTD

MltF-NTD contains two α/β subdomains, that form a central groove at their interface and that are connected by a linker region that is composed of two β-strands (Figure 1). Subdomain S1, encompassing residues 34-124 and 229-276, contains five α-helices (α1-α3, α7, α8) and two additional β-strands (β2-β3) packed against a five-stranded mixed parallel/antiparallel β-sheet (β1, β4-β6, β14). The strand order of the β-sheet is β4-β1-β5-β14-β6, with β14 following the crossover from subdomain S2 back to S1 and being the only anti-parallel β-strand in the sheet. Subdomain S2, encompassing residues 129-226 has a similar five-stranded mixed parallel/antiparallel β-sheet (β8-β12), surrounded by three α-helices (α4-α6). The topology of the central β-sheet is

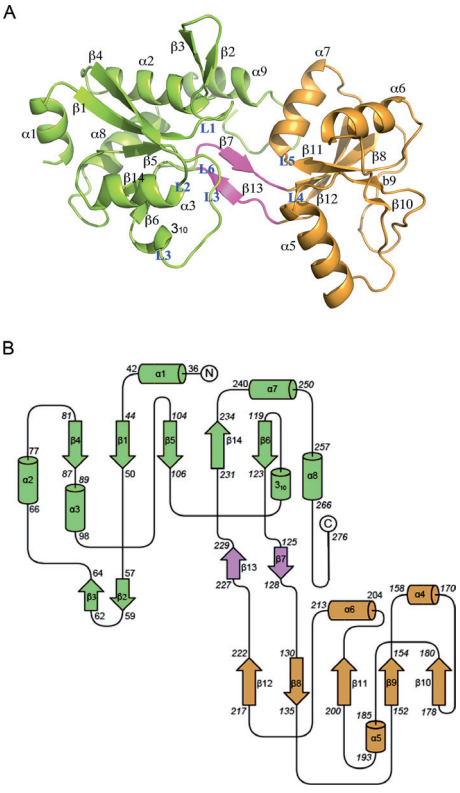

Figure 1.
Overall structure of MltF-NTD. (A) Cartoon representation of the protein backbone with subdomains S1 and S2 colored in green and orange, respectively, and the two subdomain-interconnecting strands colored in magenta. Secondary-structure elements were defined with the program DSSP (B) Topology diagram of MltF-NTD, using the same coloring scheme as in (A).

identical to that in S1, with β8 running anti-parallel with respect to the other strands and following the crossover from S1 to S2. The sheet topology of the two subdomains and the two β-strand subdomain linkers (β7 and β13) clearly define MltF-NTD as a type-2 SBP.[9] A third β-strand (β15) interacts with the two β-strand linker, but this extra strand is formed by residues of the poly-histidine tag and thus represents an artificial extension of the protein domain. The MltF-NTD structure contains no bound ligand at the inter-subdomain groove, and from the relative disposition of the two subdomains it is evident that it adopts an open conformation.

Comparison with other SBPs

A DALI search revealed a number of proteins with high structural similarity to MltF-NTD. Among the highest matches was the *E. coli* glutamine binding protein (GlnBP) [Z=19.5 with 209 of the 220 Cα atoms aligning with an RMSD

of 3.2 Å; PDB entry 1GGG],[27] the *Salmonella typhimurium* LAO-binding protein (LAOBP) [Z=19.4 with 222 of the 238 Cα atoms aligning with an RMSD of 3.4 Å; PDB entry 2LAO],[28] and the *E. coli* histidine-binding protein (HisJ) [Z=16.8 with 132 of 238 Cα atoms aligning with an RMSD of 2.1 Å; PDB entry 1HSL].[29] The identified proteins with high structural similarity also included a number of unpublished entries in the PDB, annotated as putative amino-acid-binding proteins of ABC transporters, and several eukaryotic ionotropic L-glutamate receptors and their bacterial homologs. All these close structural homologs are type-2 SBPs that belong to cluster F, following the recent classification by Berntsson *et al.*[10] The sequence identity of MltF-NTD with these type-2 SBPs is low and ranges from 15-24 %. As expected, the overall structural similarity of MltF-NTD matches best the "open-unliganded" conformations of the SBPs, i.e. the "open-unliganded" structures of GlnBP and LAOBP are ranked significantly higher than their "closed-liganded" structures.[27,28,30] The HisJ structure has only been determined in a "closed-liganded" conformation.[29] The structure of MltF-NTD, however, shows a more open conformation than most "open-unliganded" SBPs (Figure 2A). Compared to the "open" structures of GlnBP and LAOBP, the two subdomains in MltF-NTD are more widely apart and separated by an extra rotation of ~25-30 degrees.

The structural relationships become more apparent by analyzing individual structural superpositions of the subdomains of sMltF-NTD with the domains in the structures of GlnBP and LAOBP ("open-unliganded" and "closed-liganded") and HisJ ("closed-liganded" (Figure 2B). The overall folds of S1 and S2 in sMltF-NTD are highly similar to those of their counterparts in the type 2 amino-acid-binding proteins (RMSDs for the Cα-backbone overlays range from 1.9-2.5 Å), with significant differences being mainly confined to the N- and C-terminal regions. One pronounced difference is observed for the loop that connects β5 and β6 in S1 of sMltF-NTD (we named this segment the L3 loop, as explained below). Relative to the equivalent loops in the amino-acid-binding proteins, the L3 loop in MltF-NTD has shifted to a more outward position, away from the core of S1 towards the inter-subdomain cleft. The difference is remarkable, as the equivalent loop in the amino-acid-binding proteins has an important role in substrate binding and its conformation is highly conserved and relatively rigid, as judged from comparing the "open-unliganded" and "closed-liganded" structures of GlnBP and LAOBP. The conformational difference likely has consequences for the ligand-binding potential of sMltF-

NTD, as will be further discussed below. A further difference is observed for the β-strand that forms the first inter-subdomain linker in sMltF-NTD (β7). This region in MltF-NTD is equivalent to that in GlnBP, but shorter than the corresponding linker regions in LAOBP and HisJ, which contain an insertion of five amino acid residues. The second interdomain linker (β13) in sMltF-NTD is similar to that in all three amino-acid-binding proteins. In SBPs the linker regions function as flexible hinges allowing the relative movement of the two domains as rigid-bodies. Unfortunately, we lack information that could signify whether sMltF-NTD exhibits a similar flexibility leading to a closure of its inter-subdomain cleft.

Characteristics of the potential substrate-binding site in MltF-NTD

While it is evident that the overall structure of MltF-NTD is closely related to type-2 periplasmic SBPs that bind polar and charged amino acid residues, the question remains whether this domain is also a functional SBP. Amino-acid binding proteins like GlnBP, LAOBP and HisJ recognize and bind their cognate ligands in a highly similar way and use a ligand binding pocket constructed from similar segments in the polypeptide chain.[6,27-30] All type-2 amino-acid-binding proteins use a conserved arginine residue (Arg75 in GlnBP, Arg77 in LAOBP, Arg77 in HisJ), threonine/serine residue (Thr70 in GlnBP, Ser72 in LAOBP, Ser72 in HisJ) and aspartate residue (Asp157 in GlnBP, Asp161 in LAOBP, Asp161 in HisJ) to stabilize the α-ammonium and α-carboxyl groups of the amino acid ligands. The overall binding modes of the ligand side-chains are also similar, although the specific interactions are much more varied, as required to match differences in size, shape and charge. Hydrophobic interactions stabilize the stem of the side-chains, provided mainly by tyrosine, phenylalanine or leucine residues, while the head groups are stabilized via hydrogen bonding with specific protein residues and water molecules. In addition, in some of these SBPs the head group of their cognate ligand is stabilized via an interaction with a second conserved aspartate residue (Asp10 in GlnBP, Asp11 in LAOBP). To uncover whether MltF-NTD has a ligand-binding site we compared the ligand-binding regions of LAOBP, GlnBP and HisJ with the structurally equivalent regions in MltF-NTD (Figure 3). Based on the overall structural similarity with the amino-acid-binding proteins six potential ligand-binding regions can be discerned in MltF-NTD (we named these regions L1-L6), which comprise loops

and termini of secondary structure elements that face the inter-subdomain cleft. Only two of these potential ligand-binding regions show a significant sequence identity with GlnBP, LAOBP, or HisJ. These are the aforementioned β6-β6 loop (L3) and the β11-α5 turn (L5), which in the amino-acid-binding proteins contain the conserved arginine, aspartate and threonine/serine residues that bind the α-ammonium and α-carboxyl groups of the amino acid ligand. However, only the arginine residue in the L3 loop and the aspartate residue in the L5 turn are conserved in MltF-NTD (Arg125 and Asp203), while the threonine/serine residue in the L3 loop is replaced by a valine residue (Val110). Furthermore, as already mentioned, the L3 loop adopts a significantly different conformation in MltF-NTD, and in the superposition with the "ligand-closed" structures of GlnBP, LAOBP and HisJ, it actually partially overlaps with the position of the bound ligands (Figure 4). In addition, the side chain of Arg115 in MltF-NTD

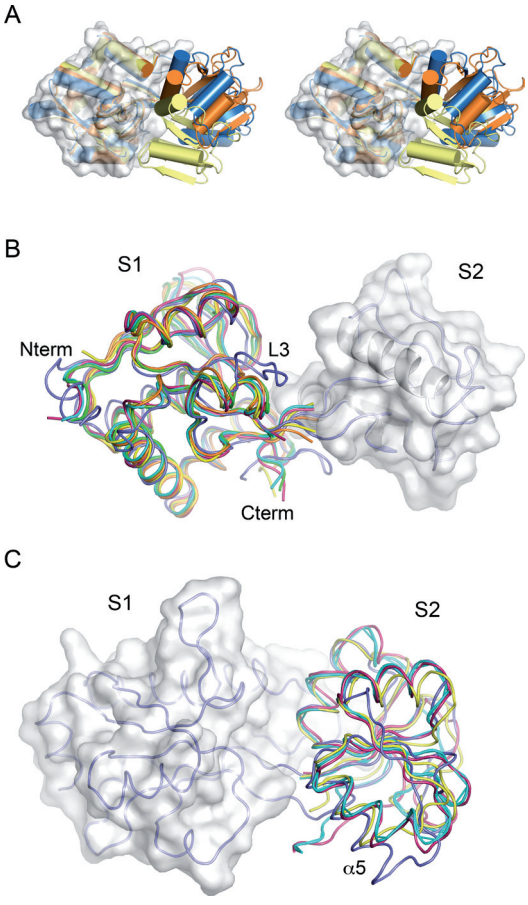

Figure 2
Comparison of MltF-NTD with type-2 amino-acid-binding proteins (A) Stereo diagram showing an overlay of the structure of MltF-NTD (dark blue) with the "open-unliganded" structure and "closed-liganded" structures of GlnBP (orange and yellow, respectively, PDB entries 1GGG and 1WDN)[27,30]. (B) Superpositions of subdomain S1 of MLtF-NTD (blue) on the equivalent domains of GlnBP (orange, "open-unliganded"; yellow, Gln-bound) LAOBP (green, "open-unliganded"; yellow, Lys-bound) and HisJ (red, His-bound) (C) Superpositions of subdomain S2 on the equivalent domains of GlnBP, LAOBP and HisJ

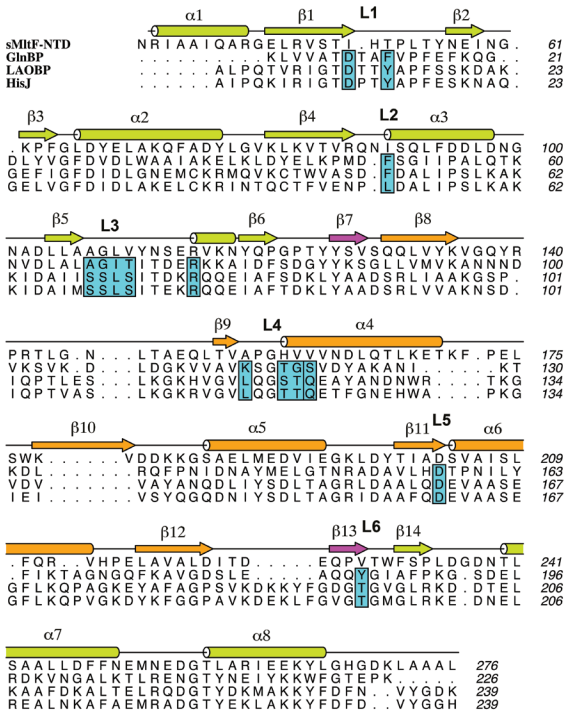

Figure 3
Structure-based multiple sequence alignment of MltF-NTD with GlnBP, LAOBP and HisJ.

points away from the potential ligand site, a difference in conformation that is stabilized by a salt-bridge interaction between Arg115 and Asp95. Finally, the conformation of the L3 loop in MltF-NTD is incompatible with a full closure of the inter-subdomain cleft to form a closed MltF-NTD structure similar as the "ligand-closed" structures of the amino-acid-binding proteins, since the L3 loop would clash with helix α5 of S2 at the subdomain interface. Thus, unless its L3 loop undergoes a dramatic conformational change upon ligand binding or upon closure of the inter-subdomain cleft, it seems unlikely that MltF-NTD binds similar amino acid ligands as GlnBP, LAOBP and HisJ. It is also possible, however, that MltF-NTD does not bind any ligand, and merely has a role as spacer-domain, to allow the C-terminal domain of the outer-membrane-bound protein to reach the peptidoglycan layer. Indeed, binding assays with different amino acid residues, small peptidoglycan-derived molecules, as well as isolated peptidoglycan have so far not been able to reveal a ligand-binding role for MltF-NTD (our unpublished data). On the other hand, the residues in the potential ligand binding regions of MltF-NTD are highly conserved in MltF homologs from different bacterial species, strongly suggesting that these regions have a

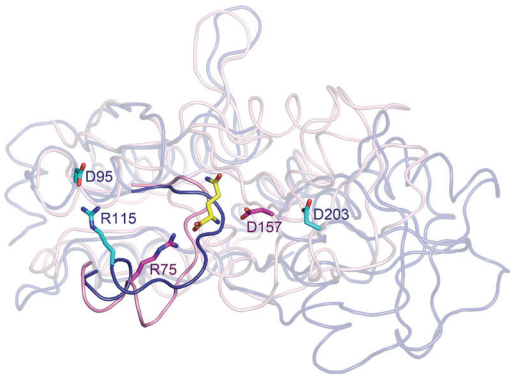

Figure 4
Different conformation of the L3 loop in MltF-NTD. superposition of the L3 loop and L5 loop in MltF-NTD (blue) and Gln-bound GlnBP (pink)

functional role (Figure 5). Unfortunately, the current structural and functional data are insufficient to clarify whether MltF-NTD is functional SBP.

CONCLUSION

In this study, we have presented the crystal structure of the SBP-like N-terminal domain of MltF, a membrane-attached lytic transglycosylase from *E. coli*, at 2.2 Å resolution. To the best of our knowledge, this is the first structure of an SBP that is present as a domain of an outer-membrane attached periplasmic protein. The structure of MltF-NTD displays close structural relationships with type-2 SBPs, in particular with the periplasmic SBPs of Gram-negative bacteria that participate in the binding and transport of polar and charged amino acid residues. In the crystal used to determine the structure of MltF-NTD the protein has an "open-unliganded" conformation, but it is unknown whether in solution the protein adopts different conformations, perhaps including those that more resemble a closed conformation. It is also unclear whether MltF-

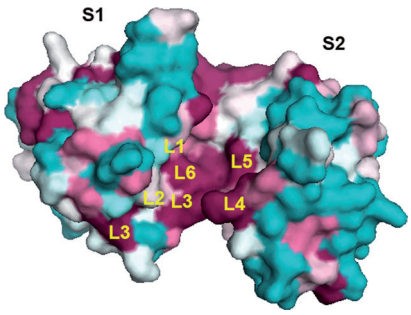

Figure 5
Residue conservation in MltF-NTD. Surface representation of MltF-NTD, colored according to residue conservation as analyzed by the program Consurf. A total of 25 MltF proteins from different bacterial species were identified (cut-off of 50% sequence identity and used for the analysis of residue conservation. The colors vary from dark purple to cyan for highly and weakly conserved residues, respectively

NTD has a ligand-binding function, but considering the differences in amino acid side chains and local backbone conformations at the inter-subdomain cleft, it seems unlikely that MltF-NTD binds similar amino acids as its SBP counterparts. Thus, the exact role of the N-terminal domain of MltF must await further biochemical and/or structural analysis.

Table 1: Summary of model refinement statistics

Space group	$P3_1$ 2 1
Cell dimensions	
a, b, c (Å)	82.4, 82.4, 75.1
α, β, γ (°)	90, 90, 120
Refinement	
Resolution (Å)	25.4-2.2 Å
No. reflections	15184
Rwork / Rfree	0.19/0.23
No. atoms	1958
Protein (MltF-NTD monomer)	1896
Water	62
B-factors (Å^2)	
Protein	58.3 Å^2
Ligand/ion	-
Water	52 Å^2
r.m.s. deviations	
Bond lengths (Å)	0.008
Bond angles (°)	1.063
Ramachandran statistics of ϕ/ψ angles (%)	
Most favored	97.9
Allowed	2.1
Molprobity score	2.0

Acknowledgements

We thank the staff of the ESRF (Grenoble) for providing facilities for diffraction measurements and for assistance. Work was supported in part by an Ubbo Emmius Bursary from the University of Groningen awarded to P.K.M.

References

1. Armstrong, N. & Gouaux, E. (2000). Mechanisms for activation and antagonism of an AMPA-sensitive glutamate receptor: crystal structures of the GluR2 ligand binding core. *Neuron* **28**, 165-81.
2. Davidson, A. L., Dassa, E., Orelle, C. & Chen, J. (2008). Structure, function, and evolution of bacterial ATP-binding cassette systems. *Microbiol Mol Biol Rev* **72**, 317-64, table of contents.
3. Felder, C. B., Graul, R. C., Lee, A. Y., Merkle, H. P. & Sadee, W. (1999). The Venus flytrap of periplasmic binding proteins: an ancient protein module present in multiple drug receptors. *AAPS PharmSci.* **1**, 1-20.
4. Gonin, S., Arnoux, P., Pierru, B., Lavergne, J., Alonso, B., Sabaty, M. & Pignol, D. (2007). Crystal structures of an Extracytoplasmic Solute Receptor from a TRAP transporter in its open and closed forms reveal a helix-swapped dimer requiring a cation for alpha-keto acid binding. *BMC Struct Biol* **7**, 11.
5. Lewis, M., Chang, G., Horton, N. C., Kercher, M. A., Pace, H. C., Schumacher, M. A., Brennan, R. G. & Lu, P. (1996). Crystal structure of the lactose operon repressor and its complexes with DNA and inducer. *Science* **271**, 1247-54.
6. Quiocho, F. A. & Ledvina, P. S. (1996). Atomic structure and specificity of bacterial periplasmic receptors for active transport and chemotaxis: variation of common themes. *Mol Microbiol* **20**, 17-25.
7. Tam, R. & Saier, M. H., Jr. (1993). Structural, functional, and evolutionary relationships among extracellular solute-binding receptors of bacteria. *Microbiol Rev* **57**, 320-46.
8. Sack, J. S., Saper, M. A. & Quiocho, F. A. (1989). Periplasmic binding protein structure and function. Refined X-ray structures of the leucine/isoleucine/valine-binding protein and its complex with leucine. *J Mol Biol* **206**, 171-91.
9. Fukami-Kobayashi, K., Tateno, Y. & Nishikawa, K. (1999). Domain dislocation: a change of core structure in periplasmic binding proteins in their evolutionary history. *J Mol Biol* **286**, 279-90.
10. Berntsson, R. P., Smits, S. H., Schmitt, L., Slotboom, D. J. & Poolman, B. (2010). A structural classification of substrate-binding proteins. *FEBS Lett* **584**, 2606-17.
11. Scheurwater, E. M. & Clarke, A. J. (2008). The C-terminal domain of Escherichia coli YfhD functions as a lytic transglycosylase. *J Biol Chem* **283**, 8363-73.
12. Scheurwater, E., Reid, C. W. & Clarke, A. J. (2008). Lytic transglycosylases: bacterial space-making autolysins. *Int J Biochem Cell Biol* **40**, 586-91.
13. Thunnissen, A. M., Rozeboom, H. J., Kalk, K. H. & Dijkstra, B. W. (1995). Structure of the 70-kDa soluble lytic transglycosylase complexed with bulgecin A. Implications for the enzymatic mechanism. *Biochemistry* **34**, 12729-37.
14. van Asselt, E. J., Dijkstra, A. J., Kalk, K. H., Takacs, B., Keck, W. & Dijkstra, B. W. (1999). Crystal structure of Escherichia coli lytic transglycosylase Slt35 reveals a lysozyme-like catalytic domain with an EF-hand. *Structure* **7**, 1167-80.
15. van Straaten, K. E., Dijkstra, B. W., Vollmer, W. & Thunnissen, A. M. W. H. (2005). Crystal structure of MltA from reveals a unique lytic transglycosylase fold. *Journal of Molecular Biology* **352**, 1068-1080.

16. Thunnissen, A. M., Isaacs, N. W. & Dijkstra, B. W. (1995). The catalytic domain of a bacterial lytic transglycosylase defines a novel class of lysozymes. *Proteins* **22**, 245-58.
17. Madoori, P. K. & Thunnissen, A. M. (2010). Purification, crystallization and preliminary X-ray diffraction analysis of the lytic transglycosylase MltF from Escherichia coli. *Acta Crystallogr Sect F Struct Biol Cryst Commun* **66**, 534-8.
18. McCoy, A. J. (2007). Solving structures of protein complexes by molecular replacement with Phaser. *Acta Crystallogr D Biol Crystallogr* **63**, 32-41.
19. Terwilliger, T. C. (2003). Improving macromolecular atomic models at moderate resolution by automated iterative model building, statistical density modification and refinement. *Acta Crystallogr D Biol Crystallogr* **59**, 1174-1182.
20. Emsley, P. & Cowtan, K. (2004). Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* **60**, 2126-32.
21. Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997). Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr D Biol Crystallogr* **53**, 240-55.
22. Adams, P. D., Afonine, P. V., Bunkoczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L. W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C. & Zwart, P. H. (2010). PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* **66**, 213-21.
23. Painter, J. & Merritt, E. A. (2006). Optimal description of a protein structure in terms of multiple groups undergoing TLS motion. *Acta Crystallogr D Biol Crystallogr* **62**, 439-50.
24. Davis, I. W., Leaver-Fay, A., Chen, V. B., Block, J. N., Kapral, G. J., Wang, X., Murray, L. W., Arendall, W. B., 3rd, Snoeyink, J., Richardson, J. S. & Richardson, D. C. (2007). MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. *Nucleic Acids Res* **35**, W375-383.
25. Waterhouse, A.M., Procter, J.B., Martin, D.M.A, Clamp, M. & Barton, G. J. (2009). Jalview Version 2 - a multiple sequence alignment editor and analysis workbench. *Bioinformatics* **25**, 1189-1191.
26. Ashkenazy H., Erez E., Martz E., Pupko T. & Ben-Tal N. (2010) ConSurf 2010: calculating evolutionary conservation in sequence and structure of proteins and nucleic acids. *Nucl. Acids Res.* **38**, W529-533.
27. Hsiao, C. D., Sun, Y. J., Rose, J. & Wang, B. C. (1996). The crystal structure of glutamine-binding protein from *Escherichia coli*. *J. Mol. Biol.* **262**, 225-242.
28. Oh, B. H., Pandit, J., Kang, C. H., Nikaido, K., Gokcen, S., Ames, G. F. & Kim, S. H. (1993). Three-dimensional structures of the periplasmic lysine/arginine/ornithine-binding protein with and without a ligand. *J Biol Chem* **268**, 11348-55.
29. Yao, N., Trakhanov, S. & Quiocho, F. A. (1994). Refined 1.89-A structure of the histidine-binding protein complexed with histidine and its relationship with many other active transport/chemosensory proteins. *Biochemistry* **33**, 4769-79.
30. Sun, Y. J., Rose, J., Wang, B. C. & Hsiao, C. D. (1998). The structure of glutamine-binding protein complexed with glutamine at 1.94 A resolution: comparisons with other amino acid binding proteins. *J Mol Biol* **278**, 219-29.

Summary

Multidrug resistance (MDR) enables cells, such as bacterial and cancer cells, to withstand the toxicity of a variety of chemically different compounds. The emergence of MDR in disease-causing bacteria has resulted in the rise of superbugs, which have developed resistance to most available antibiotics. As a result the treatment of infectious diseases has become much more difficult as it necessitates the use of last-resort drugs and even then is sometimes unsuccessful. One of the main questions that remain largely unanswered is how the MDR strains identify all these different chemical compounds. Several types of resistance mechanisms have been identified, as discussed in **Chapter 1**. Efflux pumps are membrane-bound transporters that have been proposed to be responsible for the identification and excretion of chemically different drugs preventing them from reaching their cellular targets. Very few structures of efflux pumps in the drug-bound form are available, and the structures that have been solved are of too low resolution to accurately depict and analyze the protein-drug interactions.

MDR-like efflux pumps are ubiquitous in bacteria and have roles in many metabolic pathways. In a toxic environment (e.g. due to the presence of antibiotics) the expression of efflux pumps often increases to allow cell survival. Transcriptional regulators, which are DNA binding proteins, facilitate the up-regulation of the efflux pumps. In general, in the presence of toxic compounds, transcriptional regulators respond by binding to these compounds and enhancing the transcription of efflux pumps. For example, the well-characterized local transcriptional activator BmrR of *Bacillus subtilis* and the transcriptional repressor QacR of *Staphylococcus aureus* regulate the expression of the MDR efflux pumps Bmr and QacA, respectively. In order to perform their regulatory roles, MDR-related transcriptional regulators often exist as dimers. Each subunit comprises two domains, an N-terminal DNA-binding domain and a C-terminal domain that is involved in dimerization and ligand binding. A helix-turn-helix motif located in the DNA binding domain facilitates DNA binding. Several families of MDR-related transcriptional regulators, such as TetR, MarR, MerR, AraC and PadR-like have been proposed based on commonalities in the DNA binding motif. The ligand binding domains, though they bind chemically similar ligands, vary structurally across the families and highlight the importance of studying the principles of multidrug specificity. Unlike the membrane-bound MDR transporters, transcriptional regulators are more accessible targets for structural- and drug-binding studies

as they can be easily expressed and purified in high amounts and are readily crystallizable.

The Gram-positive bacterium *Lactococcus lactis* is a non-pathogenic bacterium that is widely used in the production of fermented foods. *L. lactis* has developed an MDR phenotype upon exposure to structurally unrelated compounds such as daunomycin, Hoechst 33342, ethidium and rhodamine 6G. This MDR phenotype is due to the constitutive expression of *lmrCD* genes that encode an ATP-Binding Cassette (ABC) MDR transporter that excretes toxic compounds from the cell. The local transcriptional regulator LmrR regulates the expression of the *lmrCD* genes by responding to the same set of compounds as its cognate transporter LmrCD. By homology, LmrR belongs to PadR family of transcriptional regulators.

Structures of several drug-bound forms of well-studied transcriptional regulators, like QacR and BmrR, have been obtained providing different insights into one of the most intriguing molecular aspects of MDR, namely multidrug recognition at atomic resolution, as summarized in **Chapter 1**. So far two different models exist to explain the phenomenon of multidrug recognition: a canonical "multisite" model, in which distinct overlapping minipockets and flexible elements present in a large single drug-binding site facilitate multidrug recognition. The second model suggests a smaller and single rigid drug-binding pocket, composed of a hydrophobic slot for binding the common, aromatic cores of the target drugs, and a hydrophilic cavity for promiscuous binding of the variable drug constituents. Due to the varied nature of the drug-binding sites several more examples are required for further advancement in understanding the principles of multidrug recognition. In the first part of this thesis (**Chapters 2-3**) we describe another unique model of multidrug recognition occurring in LmrR as deduced from X-ray crystallography. This technique allows us to probe the three-dimensional structures of proteins from which we can infer molecular details on how they function.

Multidrug recognition in LmrR

To understand the drug-binding features of LmrR, we solved crystal structures of this protein in a drug-free state at 2.0 Å resolution and of drug-bound forms of LmrR with Hoechst 33342 and daunomycin, both at 2.2 Å resolution. Similar

to other MDR-related transcriptional regulators, LmrR executes its functions as a dimer. In the N-terminal DNA-binding domain, LmrR contains the canonical helix-turn-helix motif. A large central pore for drug binding, which is present at the dimer interface in LmrR, deviates from the drug-binding pockets of the previously studied QacR and BmrR regulators. The central pore in LmrR is symmetric with equal contributions from both monomers and has a pair of symmetry related tryptophan residues at the entrance. These features are unique because none of the MDR-related transcriptional regulators has a central pore at their dimer interface. The drug-binding sites of BmrR from *B. subtilis* and QacR from *S. aureus* are asymmetric and are located within a single subunit and contain several aromatic residues. In LmrR the planar ring system of the lipophilic drugs Hoechst 33342 and daunomycin is wedged between the side chains of the two tryptophans, having aromatic stacking interactions with their indole rings. There are no hydrogen bonding interactions between the drug(s) and LmrR. Binding of the drug in the central pore induces a change in the spacing between the DNA recognition helices. This change is postulated to prevent LmrR from binding to the control region of DNA resulting in the upregulation of the genes encoding the MDR transporter LmrCD. Thus by employing a pair of tryptophan residues in a single flexible drug-binding site, LmrR offers a novel mode of multidrug recognition, which deviates from the two models derived from the QacR and BmrR studies.

LmrR uses a single flexible drug-binding site for various drugs

In **Chapter 3** we describe two additional drug-bound complexes of LmrR, with riboflavin (RBF) and ethidium (ET) at 2.4 Å and 2.0 Å resolution, respectively. The overall binding modes of RBF and ET are similar to those of Hoechst 33342 and daunomycin. Drug recognition is dominated by aromatic stacking interactions with the dimer-related tryptophan indole side chains. Furthermore, these complexes provide additional evidence that planar compounds with bulky hydrophilic side chains, like those of daunomycin and riboflavin, will have less binding affinity compared to those that have hydrophobic side chains, as present in ethidium. This is because hydrophilic side chains are less well accommodated within the hydrophobic binding site than the ethyl and phenyl side chains of ET. The two additional structures presented in this chapter confirm that LmrR uses a single drug-binding site within a large, symmetrical

and hydrophobic pore in the dimer interface, and that it deploys dimer-related tryptophan residues to accomplish multidrug recognition.

As drug resistance is emerging at a faster rate than new antibiotics are discovered, there is an increasing need for the identification of new microbial drug targets. Lytic transglycosylases (LTs) may represent such a new interesting family of drug targets. These proteins are a set of autolysins found in peptidoglycan-containing eubacteria. Peptidoglycan (PG) is an important component of the cell wall of eubacteria. Its mesh-like structure allows the bacterial cells to withstand the high internal osmotic pressure.

LTs are glycosyl-transferases implicated in continuous remodeling and turnover of PG. Similar to lysozymes, LTs cleave the β-1,4 glycosidic linkage between N-acetylmuramic acid (MurNAc) and N-acetylglucosamine (GlcNAc) residues present in peptidoglycan. However, unlike lysozymes, LTs also promote the formation of an intra-residue 1,6-glycosidic bond, producing GlcNAc-anhydro-muropeptides. In *Escherichia coli* LTs are well studied. There are six outer membrane-bound (MltA, MltB, MltC, MltD, MltE, MltF) and one soluble lytic transglycosylase (Slt70). Most of the LTs contain additional non-catalytic domains for which the function is often unknown. Therefore, the specific roles of the different LTs in the PG metabolism of *E. coli* have remained unclear so far. Recently, it was reported that MltF contains an additional N-terminal domain that is homologous to periplasmic substrate-binding proteins (SBPs) of ABC transporters. The function of this N-terminal domain (MltF-NTD) is not known.

Efforts to understand the role of the N-terminal domain in the lytic transglycosylase MltF

In **Chapter 4**, crystallization experiments of a soluble form of the membrane-bound lytic transglycosylase F from *Escherichia coli* (sMltF), and of its isolated N-terminal domain (sMltF-NTD), are reported together with an analysis of their 3.5 Å and 2.2 Å resolution X-ray diffraction data. Initial attempts to elucidate the structures, using molecular replacement, of sMltF and sMltF-NTD were not successful. However, a three-wavelength MAD data set collected at 2.5 Å resolution from bromide-soaked sMltF-NTD crystals enabled us to obtain phase information and subsequently allowed the building of a partial model.

Does sMltF-NTD have ligand-binding function?

The structure of sMltF-NTD is described in **Chapter 5**, confirming its relationships with periplasmic substrate binding proteins (SBPs), in particular those that are specific for glutamine, lysine/arginine/ornithine and histidine. The N-terminal domain of MltF contains two distinct subdomains connected by a hinge. MltF-NTD has an open conformation with a ligand-free inter-subdomain cleft. In structures with a similar fold, the ligands are bound in the interdomain cleft. In those cases, a highly conserved arginine and aspartate residue in the cleft stabilize the α-ammonium and α-carboxylate groups of the ligands. Similarly, sMltF-NTD also contains such residues (Arg 125 and Asp203). However, the remaining residue composition of the interdomain cleft in sMltF differs from that of its structural homologues, and, possibly, these differences enable sMltF-NTD to have a different substrate specificity. For instance, MltF-NTD may bind the reaction product delivered by its C-terminal lysozyme-like catalytic domain, it could bind breakdown products of PG found in the periplasm, it may bind completely different compounds, or perhaps it even does not bind any ligand. These possibilities need to be tested further before a firm conclusion can be reached on the role and function of sMltF-NTD.

Using the structure of sMltF-NTD, new research can now be undertaken using experimental and computational methods to identify whether it has any role in ligand binding and subsequently in PG metabolism. Such methods may also allow us to determine whether sMltF-NTD can undergo a similar cleft opening and closing that SBPs exhibit to bind to various ligands.

Samenvatting

Multidrug resistentie (MDR) geeft cellen, zoals bacterie- en kankercellen, de mogelijkheid om weerstand te bieden aan de toxiciteit van een groot scala aan chemisch diverse stoffen. Het ontstaan van MDR in ziekteverwekkende bacteriën heeft geleid tot de opmars van superbacteriën, die resistent zijn voor de meest gebruikte antibiotica. Om een infectie van deze MDR-bacteriën te behandelen wordt er gebruikt gemaakt van "laatste mogelijkheid" antibiotica en in enkele gevallen helpen deze ook niet. Hoe MDR-stammen de chemisch diverse stoffen kunnen herkennen en hoe ze de schadelijke effecten tegen gaan blijven open vragen. Een aantal verschillende resistentie-mechanismen zijn bekend, zoals beschreven in **Hoofdstuk 1**. Efflux pompen zijn membraan gebonden transporters waarvan gedacht wordt dat ze chemisch diverse stoffen kunnen herkennen en uitscheiden voordat die stoffen schade aan de cel kunnen aanrichten. Een klein aantal kristalstructuren van efflux pompen met ligand zijn beschikbaar, maar de resolutie waarbij deze structuren zijn opgehelderd is veelal zo laag dat de interactie tussen eiwit en ligand niet goed bekeken en geanalyseerd kan worden.

MDR-achtige efflux pompen komen veel voor in bacteriën en hebben functies in verschillende metabole routes. Om in een giftige omgeving (bv. door de aanwezigheid van antibiotica) een grotere kans op overleven te hebben, verhogen cellen vaak de expressie van efflux pompen. Transcriptie regulatoren zijn DNA bindende eiwitten die de expressie van eiwitten reguleren door te reageren op omgevingsfactoren. In het geval van toxische stoffen zijn er transcriptie regulatoren die deze stoffen binden en zorgen voor een verhoogde transcriptie van efflux pompen. Enkele voorbeelden hiervan zijn de goed gekarakteriseerde transcriptie activator BmrR uit *Bacillus subtilis* en de transcriptie repressor QacR uit *Staphylococcus aureus,* die respectievelijk de expressie van de MDR efflux pompen Bmr en QacA reguleren. MDR-gerelateerde transcriptie regulatoren komen vaak voor als dimeren, wat hun activiteit ten goede komt. Elke subunit bestaat uit twee domeinen, een DNA-bindend domein en een C-terminaal domein die voor de dimerisatie en ligand binding zorgt. Het DNA-bindend domein heeft de bekende 'helix-turn-helix' vouwing en de klassificatie van verschillende MDR-gerelateerde transcriptie regulators in families, zoals TetR, MarR, MerR, AraC en PadR, is gebaseerd op overeenkomsten in aminozuurvolgorde van dit domein. Voor het ligand-bindend domein geldt een ander verhaal. Deze is namelijk structureel heel

divers voor de verschillende MDR-gerelateerde transcriptie regulatoren, terwijl deze eiwitten vaak wel overeenkomstige liganden binden. Dit zijn dus interessante domeinen om de basis van multidrug specificiteit te bestuderen. In tegenstelling tot de membraangebonden MDR transporters zijn de transcriptie regulatoren voor structuur- en drugbindingsstudies makkelijker beschikbare targets, omdat de expressie en zuivering van de regulatoren eenvoudiger is en ze daarnaast makkelijker kristalliseren.

De Gram-positieve bacterie *Lactococcus lactis* is niet pathogeen en wordt zeer veel gebruikt bij het produceren van gefermenteerd voedsel. *L. lactis* heeft een MDR-fenotype als het blootgesteld wordt aan stoffen zoals daunomycine, Hoechst 33342, ethidium en rhodamine 6G, die alle chemisch gezien van elkaar verschillen. Dit fenotype ontstaat door de constante expressie van de zogenaamde *lmrCD* genen, die samen coderen voor een ATP-Bindende Cassette (ABC) MDR transporter. De locale transcriptie regulator LmrR reguleert de expressie van de *lmrCD* genen door te reageren op dezelfde set stoffen die de LmrCD transporter uit de cel pompt. Homologisch gezien behoort de LmrR tot de PadR familie van transcriptie regulatoren.

De structuren van drug-gebonden vormen van de transcriptie regulatoren, zoals QacR en BmrR, geven inzicht in één van de meer intrigerende moleculaire aspecten van MDR, namelijk hoe de chemisch verschillende stoffen herkend en gebonden worden op een atomair niveau (**Hoofdstuk 1).** Tot op heden zijn er twee modellen die de herkenning van verschillende stoffen verklaren. Als eerste het 'multi-site' model, waarin gedefinieerde overlappende minipockets en flexibele elementen samen een grote drug-bindingsplaats vormen die multidrug herkenning mogelijk maakt. Het tweede model gaat uit van een kleinere en minder flexibele bindingsplaats, bestaande uit een hydrofobe gleuf waarin aromatische groepen van een target-drug kunnen binden, met daar omheen enkele hydrofiele holtes waarin de variabele groepen die aan de aromatische groep zitten terecht kunnen. Omdat er waarschijnlijk een grote variatie bestaat in drug-bindingsplaatsen, zijn er meer structuren nodig om een beter beeld te krijgen van multidrug herkenning. In het eerste deel van dit proefschrift (**Hoofdstukken 2 en 3)** beschrijven we het model voor multidrug herkenning zoals dat voorkomt in LmrR, gebaseerd op kristallografische

data, en we tonen aan dat dit model significant verschilt van de twee eerder beschreven modellen.

Multidrug herkenning in LmrR

Om te begrijpen hoe LmrR drugs bindt, hebben we meerdere kristalstructuren opgehelderd. LmrR in ongebonden toestand is opgehelderd bij een resolutie van 2.0 Å, terwijl LmrR met Hoechst 33342 en daunomycine beide zijn opgehelderd bij een resolutie van 2.2 Å. Net als andere MDR transcriptie regulatoren moet ook LmrR dimeriseren om functioneel te zijn. Het N-terminale domein van LmrR heeft een 'helix-turn-helix' motief. Een grote en centrale drug-bindende porie, gelegen aan het dimeer raakvlak van LmrR, verschilt van de drug-bindende holtes van de QacR en BmrR regulatoren. De centrale porie in LmrR is symmetrisch en beide monomeren dragen evenveel bij aan deze porie. Verder zit er bij de ingang van de porie een symmetrie gerelateerd paar tryptofanen. Deze structuur eigenschappen komen niet voor bij de andere MDR-gerelateerde transcriptie regulatoren, wat LmrR uniek maakt. De drug bindende holtes van BmrR en QacR zijn asymmetrisch en liggen in een monomeer en bevatten verder meerdere aromatische residuen. In LmrR wordt het vlakke ringsysteem van Hoechst 33342 and daunomycine vast geklemd tussen de zijketens van de twee tryptofanen door middel van aromatische stacking. Er worden geen waterstofbruggen gevormd tussen de drugs en LmrR. Door de binding van een drug tussen de tryptofaan residuen ontstaat er een verandering in de afstand tussen de DNA bindende helices. Hoogstwaarschijnlijk kan LmrR hierdoor niet meer binden aan het DNA waardoor de genen coderend voor de MDR transporter LmrCD tot expressie worden gebracht.

LmrR gebruikt een enkele drugbindingsplaats voor multidrug-binding

In **hoofdstuk 3** beschrijven we nog twee structuren van LmrR in complex met een drug, namelijk met riboflavine (RBF), opgehelderd bij een resolutie van 2.4 Å, en met ethidium (ET), opgehelderd bij 2.0 Å. RBF en ET worden gebonden, net als Hoechst 33342 en daunomycine, door middel van aromatische stacking met de indool groepen van de twee tryptofanen. Verder ondersteunen deze

twee nieuwe structuren het feit dat vlakke stoffen met hydrofiele zijketens, zoals daunomycine en RBF een lagere affiniteit voor LmrR hebben dan stoffen met hydrofobe zijketens, zoals ET. Dit komt voornamelijk doordat de hydrofiele zijketens minder goed passen in de hydrofobe bindingsplaats. De twee structuren uit dit hoofdstuk bevestigen dat LmrR maar één bindingsplaats gebruikt voor multidrug-herkenning en multidrug-binding

De huidige antibiotica worden in een steeds hoger tempo onbruikbaar door de opkomst van resistente microben. Er is daarom een grote noodzaak antibiotica te ontwikkelen die gericht zijn op nieuwe targets. Eén van deze targets zou de familie van lytische transglycosylases (LTs) kunnen zijn, die de peptidoglycaan laag (PG) in de celwand van bacteriën continu herstructureren. Het PG polymeer in de celwand zorgt ervoor dat bacteriën de hoge interne osmotische druk het hoofd kunnen bieden. De LTs verbreken, net als lysozym, de β-1,4-suikerbinding tussen N-acetylmuraminezuur (MurNAc) en N-acetylglucosamine (GlcNAc), die aanwezig zijn in het PG. Ze verschillen echter van lysozym omdat ze de vorming van een intra-residuele 1,6-suikerbinding katalyseren, wat resulteert in de vorming van GlcNAc-anhydro-muropeptides. De LTs uit *Escherichia coli* zijn goed bestudeerd; er zijn zes membraangebonden LTs (MltA, MltB, MltC, MltD, MltE, MltF) en één ongebonden LT (Slt70). De meeste LTs hebben, naast hun katalytische domein, één of meerdere andere domeinen waarvan de functie nog niet bekend is. Daarom zijn de specifieke rollen van deze *E. coli* LTs in het PG metabolisme nog onduidelijk. Uit een recente studie blijkt dat MltF een N-terminaal domein heeft dat homoloog is aan periplasmatische substraatbindende eiwitten (SBPs) van ABC transportereiwitten, maar de functie van dit N-terminale domein is echter niet bekend.

Pogingen om de rol van het N-terminale domein van de lytische transglycosylase MltF te begrijpen

In **Hoofdstuk 4** staat beschreven hoe een oplosbare vorm van MltF (sMltF), en zijn N-terminale domein (sMltF-NTD), zijn gekristalliseerd, en hoe deze kristallen zijn gebruikt voor röntgendiffractie experimenten. De diffractie van de sMltF kristallen had een maximaal oplossend vermogen van 3.5 Å en van de sMltF-NTD kristallen van 2.2 Å, maar pogingen om fase informatie te verkrijgen met behulp van de 'molecular replacement' methode wilden niet

lukken. Met behulp van een andere methode, namelijk "multiple-wavelength anomalous diffraction" en gebruikmakend van in een bromide oplossing gedrenkte kristallen, is het ons toch gelukt om de structuur op te helderen van sMltF-NTD.

Heeft sMltF-NTD een ligandbindende functie?

De kristalstructuur van sMltF-NTD is beschreven in **Hoofdstuk 5** en hieruit kan opgemaakt worden dat er een duidelijke relatie is met SBPs, in het bijzonder met SBPs die specifiek zijn voor glutamine, lysine/arginine/ornithine en histidine. sMltF-NTD bestaat uit twee subdomeinen die met elkaar verbonden zijn door een scharnier. In de structuur heeft sMltF-NTD een open conformatie met een ligandloze groeve tussen de twee subdomeinen. In de gerelateerde SBPs zorgt de groeve voor de binding van een ligand. In dergelijke gevallen wordt het ligand gecoördineerd door een geconserveerde arginine en asparaginezuur residue. Deze residuen komen ook voor in het sMltF-NTD (Arg125 en Asp203), maar de meeste andere aminozuren in de groeve verschillen van die in de homologe SBPs. Deze verschillen zouden erop kunnen wijzen dat sMltF-NTD een andere substraatspecificiteit heeft, en bijvoorbeeld PG-gerelateerde liganden bindt zoals het reactieproduct van het C-terminale katalytische domein of afbraakproducten van PG die vrij voorkomen in het periplasma. Maar dit alles moet nog verder onderzocht worden voordat er conclusies getrokken kunnen worden.

Nu de kristalstructuur van sMltF-NTD bekend is, kunnen er nieuwe experimenten bedacht worden die mogelijk aantonen welke liganden dit domein kan binden en of zo'n ligandbindingsfunctie een rol speelt in het PG-metabolisme. Verder zou het interessant zijn om te zien of sMltF-NTD, net als andere SBPs, open en dicht gaat als het een ligand bindt.

Acknowledgements

My journey from India to Netherlands to obtain the most difficult degree of my life was a memorable event. As the precious goals cannot be achieved alone, there are several people who helped me to surpass different levels of happiness as well as painful moments. I owe my gratitude to all those who contributed to this unimaginable journey.